DOCUMENT RESUME

ED 292 054 CS 009 057

AUTHOR Ryan, Joseph M.; Rowls, Michael D.
TITLE An Evaluation of Students' Reading Achievement in the South Carolina Basic Skills Assessment Program 1981-1985. Final Report.
INSTITUTION South Carolina Univ., Columbia. Coll. of Education.
PUB DATE Jan 88
NOTE 248p.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC10 Plus Postage.
DESCRIPTORS *Basic Skills; Elementary Education; Inferences; Mathematics Achievement; *Program Evaluation; *Reading Achievement; Reading Comprehension; Reading Difficulties; Reading Improvement; Reading Instruction; Reading Research; Reading Tests
IDENTIFIERS *South Carolina Basic Skills Assessment Program

ABSTRACT

In order to provide a comprehensive view of the cumulative effect of the South Carolina Basic Skills Assessment Program (BSAP) on students' learning, a study examined reading and mathematics achievement by all students in the program over a 5-year period (1981-1985), as measured by the testing portion of the BSAP. Subjects were students in grades 1, 2, 3, 6, and 8. The study was organized around four research questions: (1) whether there was progress in South Carolina students' overall performance in reading and mathematics; (2) whether those who achieved both the reading and mathematics standards in 1985 exceeded the standards by as much as the corresponding students in 1981, and, whether those who reached neither standard in 1985 were any closer to reaching the reading standard than the corresponding students in 1981; (3) with which areas of reading performance were students continuing to have difficulty; and (4) what recommendations could be made to educators. Analyses revealed that students made considerable progress in all grades in both reading and mathematics. In addition, the high achieving students in 1985 were achieving at higher levels than their counterparts in 1981; while the low achieving students in 1985 were achieving at higher levels than their counterparts in 1981. Investigation of the third question revealed that, in general, all students were less capable on reading tasks that require inferential comprehension and did better on literal comprehension tasks. Finally, the findings suggest that educators should focus more effort on teaching inferential comprehension skills. (ARH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED292054

An Evaluation of Students' Reading Achievement in the South Carolina Basic Skills Assessment Program 1981-1985

Joseph M. Ryan
Department of Educational Psychology

Michael D. Rowls
Department of Instruction and Teacher Education

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Joseph M. Ryan

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

College of Education
The University of South Carolina

This research was supported by funds from the South Carolina Department of Education

BEST COPY AVAILABLE

AN EVALUATION OF STUDENTS' READING ACHIEVEMENT IN THE SOUTH CAROLINA BASIC SKILLS ASSESSMENT PROGRAM: 1981-1985

FINAL REPORT

Joseph M. Ryan
Department of Educational Psychology

Michael D. Rowls
Department of Instruction & Teacher Education

University of South Carolina
January, 1988

PREFACE

Student achievement in reading has been assessed using the tests of the South Carolina Basic Skills Assessment Program (BSAP) since 1981 in grades 1, 2, 3, 6 and 8. Yearly reports provide public information about students' achievement but these annual reports are not designed to provide a comprehensive view of student achievement over extended periods of time.

The results of a comprehensive study examining students' BSAP reading performance from 1981 to 1985 are described in this report. The report is organized around four research questions. Following an introductory section, the first major research question deals with South Carolina students' overall performance in reading and mathematics. Analyses reveal considerable student progress in all grades across the five year period in both reading and mathematics.

The second research question examines the performance of students in the high and low ranges of the reading test scales. Analyses of these students reveal that the relatively high achieving students in 1985 are achieving at higher levels than the relatively high scoring students in 1981. Also, the relatively low achieving students in 1985 are achieving at higher levels than the relatively low achieving students in 1981. In short, students at all achievement levels have evidenced progress in reading over the five year period.

The third research question attempts to identify any areas of reading performance in which students continue to have difficulty, despite the evidence of overall progress as revealed by the answers to the first two research questions. The analyses of students' deficiencies are designed

specifically for classroom teachers, school principals, reading coordinators and all other educators who can influence school curriculum and instruction. The description of students' deficiency areas in the BSAP reading objectives and subskills is described in considerable detail with numerous examples on pages 31 to 194. The analyses of students' weaknesses in reading are examined for grades 1, 2 and 3, and then for grades 6 and 8. A concise description of each subskill for each objective precedes the analysis of students' deficiencies on the subskills. These subskill descriptions reflect the critical information from Teaching and Testing Our Basic Skills. This "mini T&T" is designed to help teachers recall what is assessed by the various BSAP subskills.

In general, all students are clearly less capable on reading tasks that require inferential comprehension and do better on literal comprehension tasks. In grades 1, 2 and 3, students' reading difficulties are related to weaknesses in inferential comprehension in general, lack of familiarity with the reading tasks required by certain of the BSAP subskills, over-reliance on key and signal word cueing, pronoun substitution, and inter-sentence comprehension tasks. In grades 6 and 8, students' reading difficulties are likewise related to weaknesses in inferential comprehension in general, over-reliance on key and signal word cueing, inter-paragraph combining of information and inferences, confusion concerning superordinate and subordinate text relationships, and confusion with unrestricted or divergent reading selections as opposed to restricted or convergent reading selections.

The final section of the report is "Summary, Conclusions and Recommendations". A condensed version of the entire report is presented in this section. This section is provided for those who may not have an

opportunity to read the entire report. A variety of recommendations are suggested in this final section. Two general and important recommendations are: 1) the reallocation of instructional time to focus more effort on teaching inferential comprehension skills, and 2) more instruction of a direct nature aimed at improving inferential comprehension skills.

ACKNOWLEDGEMENTS

This report culminates a project which has taken more than three years to complete and has received the support, assistance and encouragement of many educators from all over the state of South Carolina. It is impossible to recognize all the people who have helped in this project, but we would like to offer our appreciation to some of the people who provided invaluable professional insight and personal support.

We would like to thank all our colleagues and friends at the South Carolina Department of Education and extend special appreciation to Joel Taylor, Susan Smith, and Mary Kelly in the Office of General Education; and Paul Sandifer, Mimi Todd, and Lyn Z. Mueller in the Office of Research.

Ms. Lily Chang of the Department of Educational Psychology, University of South Carolina, conducted the statistical analyses on which the study is based. We were very fortunate to have Ms. Chang work on the project with us and we would like to thank her for her outstanding and diligent effort.

We would like to thank Annette Gordon for her careful and conscientious typing of the many versions of this report.

This project was supported by funds from the South Carolina Department of Education. We want to acknowledge with special thanks the support of Dr. Charlie G. Williams, State Superintendent of Education.

JMR MDR

Table of Contents

PREFACE---------- i

ACKNOWLEDGEMENTS---------- iv

LIST OF FIGURES---------- viii

LIST OF TABLES---------- ix

INTRODUCTION

Research Questions---------- 1

Students Studied---------- 2

General Procedures---------- 3

A Conceptual Model for Teaching and Testing Reading Skills---------- 5

A Description of the BSAP Tests---------- 9

RESULTS

Research Question 1---------- 14

Overall Students' Performance on BSAP Examinations---------- 14

Discussion and Interpretation---------- 19

RESULTS

Research Question 2---------- 18

A More Detailed Examination of Students' Overall Performance---------- 18

Discussion and Interpretation---------- 30

RESULTS

Research Question 3---------- 31

Analyzing Students' Deficiencies in Reading---------- 30

Students' Performance on the Reading Objectives---------- 31

Trends Across Grades---------- 40

STUDENTS' PERFORMANCE ON THE READING SUBSKILLS AND ITEMS---------- 42

Grades 1, 2, 3

Decoding and Word Meaning---------- 44

Details---------- 57

Main Idea---------- 65

Reference Usage---------- 82

Inference---------- 92

Analysis of Literature----------106

Grades 6, 8

Decoding and Word Meaning----------117

Details----------125

Main Idea----------131

Reference Usage----------148

Inference----------158

Analysis of Literature----------180

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS----------195

Question 1----------196

Question 2----------198

Question 3----------200

Grades 1, 2, 3 ----------206

Summary of 1, 2, 3----------214

Grades 6, 8 ----------217

Summary of 6, 8----------228

Question 4----------232

LIST OF FIGURES

1. Figure 1. Conceptual Model for Teaching and Testing Reading Skills---------- 6

2. Figure 2. BSAP Reading Objectives and Subskills---------- 10

3. Figure 3. Illustration of Classification of Students by S.E.M.'s---------- 20

4. Figure 4. Average Percentage of Grade 1 Students Reaching and Not Reaching Standard Who Correctly Answer Items in the Six BSAP Objectives---------- 33

5. Figure 5. Average Percentage of Grade 2 Students Reaching and Not Reaching Standard Who Correctly Answer Items in the Six BSAP Objectives---------- 34

6. Figure 6. Average Percentage of Grade 3 Students Reaching and Not Reaching Standard Who Correctly Answer Items in the Six BSAP Objectives---------- 35

7. Figure 7. Average Percentage of Grade 6 Students Reaching and Not Reaching Standard Who Correctly Answer Items in the Six BSAP Objectives---------- 36

8. Figure 8. Average Percentage of Grade 8 Students Reaching and Not Reaching Standard Who Correctly Answer Items in the Six BSAP Objectives---------- 37

9. Figure 9. Schematic diagram of difficult IN 2 Item, Grade 8---------- 173

LIST OF TABLES

1. Number (and Percentage) of Students Meeting and Failing to Meet BSAP Standards in Reading and Mathematics---------- 15

2. Increased and Decreased Percentage of Students in The +R+M and -R-M Groups from 1981 to 1982---------- 16

3. Distribution of Students by Group Membership on BSAP Reading: Grade 1---------- 23

4. Distribution of Students by Group Membership on BSAP Reading: Grade 2---------- 24

5. Distribution of Students by Group Membership on BSAP Reading: Grade 3---------- 25

6. Distribution of Students by Group Membership on BSAP Reading: Grade 6---------- 26

7. Distribution of Students by Group Membership on BSAP Reading: Grade 8---------- 27

8. 1981-1985 Changes in the Percentage of Students Passing or Failing the Reading Standard by Greater Than Two S.E.M.'s---------- 29

9. Average Item Difficulty and Item Discrimination for First Grade +R and -R Students on Decoding and Word Meaning Subskills---------- 47

10. Average Item Difficulty and Item Discrimination for Second Grade +R and -R Students on Decoding and Word Meaning Subskills---------- 49

11. Average Item Difficulty and Item Discrimination for Third Grade +R and -R Students on Decoding and Word Meaning Subskills---------- 53

12. Average Item Difficulty and Item Discrimination for First Grade +R and -R Students on Main Idea Subskills---------- 66

13. Average Item Difficulty and Item Discrimination for Second Grade +R and -R Students on Main Idea Subskills---------- 70

AN EVALUATION OF STUDENTS' READING ACHIEVEMENT IN THE SOUTH CAROLINA BASIC SKILLS ASSESSMENT PROGRAM FROM 1981 TO 1985

Introduction

Well over a million students have taken tests as part of the South Carolina Basic Skills Assessment Program (BSAP) since the testing portion of the program began in 1981. During this time annual reports of students' performance have been provided to parents, teachers, school administrators, and the public at large. These yearly reports have been useful in describing students' annual performance relative to the program's instructional objectives. Because these reports only cover one year, however, they do not provide a comprehensive view of the cumulative effect of the BSAP program on students' learning. This report describes the results of a study which examines students' BSAP test performance in 1981, 1982, 1983, 1984, and 1985.

Research Questions

The study examines general achievement in mathematics and reading and focuses on specific areas of student performance in reading. The study was designed to answer the following four questions:

Question 1

Have students in South Carolina made progress toward achieving the grade level standards in mathematics and reading from 1981 to 1985?

Question 2

Do the students who achieve both the reading and mathematics standards in 1985 reach the reading standard by as much as the corresponding stu-

dents in 1981? And, are the students who reach neither the mathematics nor reading standards in 1985 any closer to reaching the reading standard than the corresponding students in 1981?

Question 3

Are there specific areas of reading in which students across the state have deficiencies? If so, what are the deficiency areas?

Question 4

What actions might be taken to improve student achievement relative to the BSAP reading instructional objectives?

Students Studied

Only students in grades 1, 2, 3, 6 and 8 have been tested since 1981, hence, only these grades were examined in the study. The students are subdivided into four groups in various parts of the study in order to answer the four research questions. The composition of the four groups is as follows:

1) Students who reach the grade level standards both in reading and in mathematics. These students are referred to as +R+M students.
2) Students who do not reach the grade level standard in reading but reach the standard in mathematics. These students are referred to as -R+M students.
3) Students who reach the grade level standard in reading but not in mathematics. These students are referred to as +R-M students.
4) Students who reach neither the reading nor the mathematics grade level standard. These students are referred to as the -R-M students.

In certain analyses, the +R+M and +R-M students are combined to constitute the group of all students who reach the grade level standard in reading. These students are referred to as the +R students. In addition, the -R+M and -R-M students are combined for certain analyses to constitute the group of all students who do not reach the grade level standard in reading. These students are referred to as the -R students.

As a technical note it is useful to point out here that the analyses performed in this study were conducted on the entire population of students who took the BSAP tests at the various grades from 1981 to 1985. This use of population data makes the application of inferential statistics unnecessary and inappropriate. Inferential statistics are applicable when samples are drawn from a population and are used as the basis for making inferences about characteristics of the population. Since population data are used throughout to answer the research questions, no inferential techniques are needed in this study.

General Procedures

A general description of the procedures used to investigate each of the four research questions is provided in this section. More specific details of the procedures will be described in the results section for each question.

Research Question 1 was examined by comparing changes in the proportion of students in the +R+M and -R-M groups from 1981 to 1985 to determine whether there has been improved achievement in both mathematics and reading.

Research Question 2 was examined in two ways. First, students in the +R+M group in 1985 were compared to students in the +R+M group in 1981 to determine whether the +R+M students in 1985 have a higher level of reading achievement than the +R+M students in 1981. Second, students in the -R-M

group for 1985 were compared to students in the -R-M group in 1981 to determine whether the -R-M students in 1985 are closer to the reading standard than the -R-M students in 1981.

Research Question 3 was examined by comparing the performance of students in various groups (+R+M, -R+M, +R-M and -R-M) and combinations of groups on specific BSAP objectives, subskills and items to identify objectives, subskills, or items that are conspicuously more difficult for the various types of students. A more complete description of procedures used to examine Research Question 3 will be provided when the results of the analyses are described.

Research Question 4 was examined using a conceptual model for teaching and testing reading skills. This model displays the relationships among curriculum, instruction, characteristics of teachers and students, students' achievement, and the use of information about student achievement to facilitate further learning. This conceptual model will be described shortly in more detail.

A description of the conceptual model for teaching and testing reading skills and a description of the BSAP tests follows this introduction. The results of the study are then presented for the four research questions. In the final section of the report, findings of the study are summarized and conclusions and recommendations are offered.

A Conceptual Model For Teaching and Testing Reading Skills

As the reader proceeds through this report, it will be useful to have in mind the conceptual model used to organize the research and to guide the interpretations of the results of the study. The model is pictured in Figure 1 on the next page and has three major components. The first component represents four factors which affect student learning: curriculum characteristics, instructional characteristics, student characteristics and teacher characteristics. These four factors operate separately and in combination with each other to influence student achievement. In the second component, student reading achievement which includes reading achievement in general and reading achievement related specifically to BSAP instructional objectives is portrayed. Third, the connection between reading achievement and the factors which affect further student learning is represented in the model as "Information about Achievement."

These three components work together so that the factors which affect student learning (e.g. curriculum, instruction, teachers and students) lead to achievement. Then, information about students' current level of achievement is used as the basis for modifying and recombining the factors that affect student learning to further influence student achievement. The process depicted in the model presumes a continuous "interactive loop" among the three components in the model with continuous progress in student achievement as the targeted goal.

The most important component of the model is the set of factors which affect students' learning. These factors include characteristics of the curriculum, instruction, teachers, and students. These four major factors

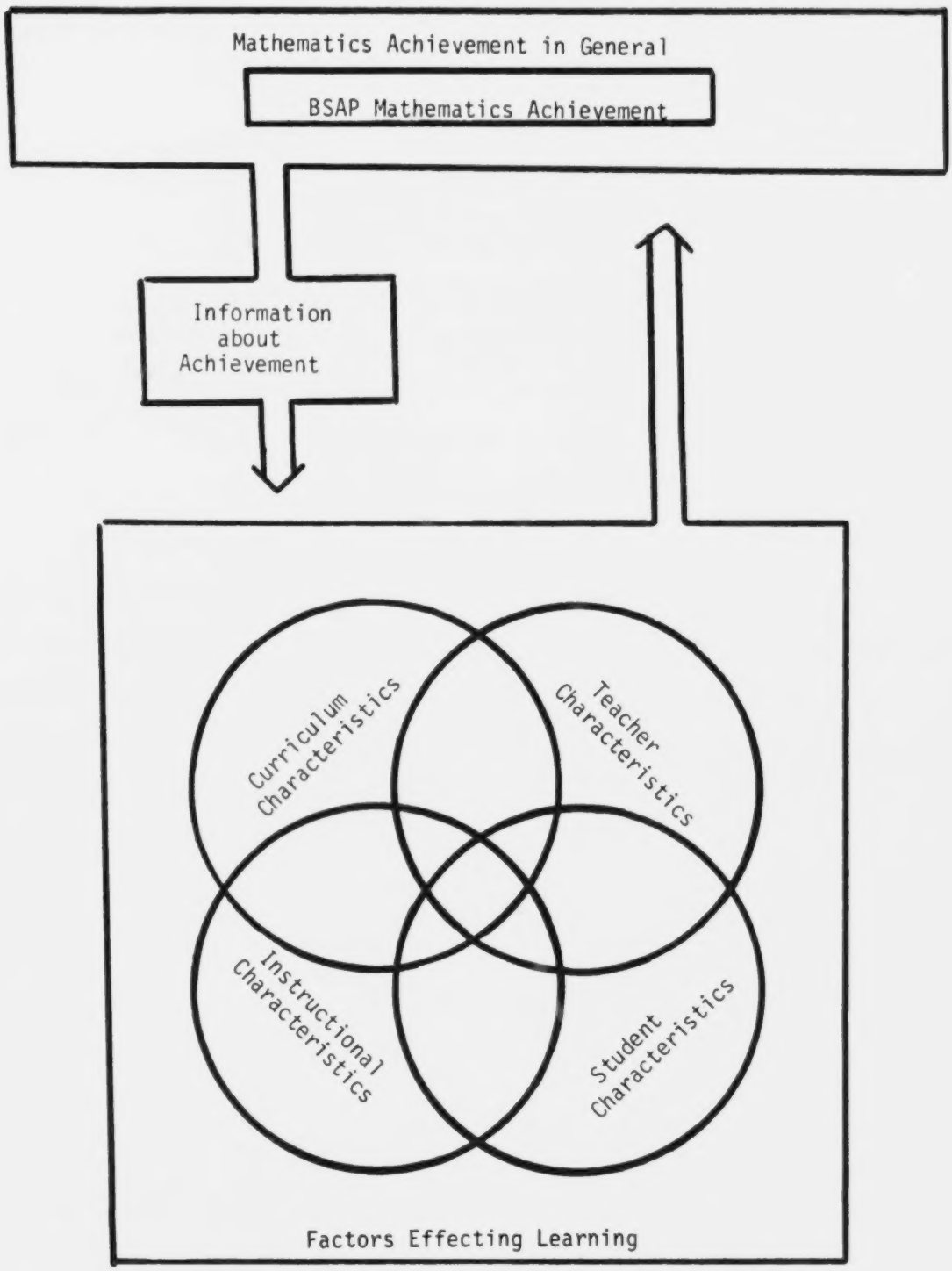


Figure 1. Conceptual Model for Teaching and Testing Mathematics Skills

interact and overlap as depicted in the model. The interactive nature of these key factors is meant to reflect the complexity of the learning environment by making it clear that certain teachers may be more effective with particular types of students using particular types of selected curriculum materials and instructional strategies. The complex interactive nature of the learning process leading to improved achievement is made explicit in the model in order to point out the inadequacy of remediation efforts which focus on only one salient element or factor of the learning process. Remediation of the reading deficiencies described in response to Research Question #3 may require the coordination of all factors which affect student learning.

The four major factors influencing student learning operate in the larger context of home, school and community environments. Characteristics of these environments clearly have an effect on student learning but are not discussed explicitly in the model since they are not easily alterable characteristics of the school environment.

In the model, the arrow from the component involving the factors affecting student learning is aimed at reading achievement to reflect the notion that such factors can indeed improve the level of student achievement in reading. Two levels of reading achievement are depicted: reading achievement in general and BSAP reading achievement in particular. Reading achievement is depicted in this two-level fashion to emphasize the fact that the BSAP Reading objectives are only a subset of school reading objectives in general. The BSAP Reading objectives were not designed as a complete and exhaustive focus for reading or learning from reading. The BSAP Reading objectives define the basic reading skills that educators hope all students will obtain at some point, but instruction should not be limited

to these skills alone.

On the left side of Figure 1 an arrow representing "Information about Achievement" proceeds from reading achievement toward the component depicting the factors which affect student learning. This "Information" component of the model is designed to represent two activities. First, information about students' current level of reading achievement must be collected through the process of assessment. The BSAP assessment is conducted annually at certain grades for reading skills specified as BSAP objectives. Secondly, assessment information must be disseminated to educators who can use the information to facilitate further student learning. This feedback typically has taken various forms: student score reports, district reports and the BSAP annual statewide report describing individual and collective student achievement in terms of the total test score, the BSAP objectives and the item response summary. This feedback component also includes other sources of information such as the results of this study and locally developed diagnostic tests which reflect students' performance relative to the BSAP objectives.

The current study began by examining the "BSAP Reading Achievement" component of the model in an effort to provide answers to research questions 1, 2, and 3. The researchers next used this information to respond to research question 4: What action might be taken to improve students achievement relative to the BSAP objectives? The answers to this research question are organized around the various components of the conceptual model and are described in detail when responses to research question 4 are presented.

A Description Of The BSAP Tests

The BSAP Reading Tests for grades 1, 2, 3, 6 and 8 assess six objectives: Decoding and Word Meaning, Details, Main Idea, Reference Usage, Inference, and Analysis of Literature. The number of subskills within each objective varies and not all subskills are tested at each grade. The subskills for each objective are shown in the chart contained in Figure 2 on the next page. The chart indicates the grades at which each subskill is tested. In this report, subskills are referred to by the first two letters of the objective and the number of the subskill (e.g., MI 2 refers to the second subskill within the Main Idea Objective).

Each BSAP Reading Test is composed of a total of 36 items, six items assessing each of the six objectives. The number of items on a test assessing the various subskills varies from year to year.

Grade level standards for all BSAP tests were set in terms of the number of items a student must answer correctly on the entire test. Standards were not set at the level of individual objectives. However, the overall standard is used to estimate a corresponding standard for each objective.

The five BSAP tests from 1981 to 1985 for any one grade are all on a common scale and the standard represents equivalent performance for each of the five years within each grade. Thus, an increase in the proportion of students meeting the first grade standard between 1981 and 1985 represents increased first grade achievement because the first grade standard is consistent and comparable across the five years.

DW: Decoding and Word Meaning

DW 1: Sight Recognition (tested at all grades)
DW 2: Phonetic Decoding (tested at grades 1,2,3)
DW 3: Contextual Word Meaning (tested at grades 2,3,6,8)
DW 4: Structural Word Meaning (tested at grades 6,8)

DE: Details

No subskills (tested at all grades)

MI: Main Idea

MI 1: Restated Main Idea (tested at grades 1,2,3)
MI 2: Paraphrased Main Idea (tested at all grades)
MI 3: Inferred Main Idea (tested at all grades)

RE: Reference Usage

RE 1: Selecting a Reference Source (tested at all grades)
RE 2: Using a Reference Source (tested at all grades)

IN: Inference

IN 1: Making Comparisons (tested at all grades)
IN 2: Determining Cause and Effect (tested at all grades)
IN 3: Drawing Conclusions (tested at all grades)
IN 4: Predicting Outcomes (tested at all grades)

AL: Analysis of Literature

AL 1: Nature of Information (tested at all grades)
AL 2: Structural Elements (tested at all grades)
AL 3: Rhetorical Devices (tested at all grades)
AL 4: Critical Analysis (tested at grade 8)

Figure 2. BSAP Reading Objectives and Subskills

The tests and standards are not equivalent across grades, however. Comparing performance across grades is not appropriate or correct. Thus, the proportion of students reaching the standard in grade 1 cannot be meaningfully compared to the proportion of students reaching the standard at grade 2 (or any other grade) since neither the student scores nor the grade level standards are equivalent across the grades.

When examining the results of this study, it is critical to remember that the BSAP examinations are not minimum competency tests. If they were, it might be reasonable to expect a large proportion or even all the students to pass the examinations. BSAP examinations assess end-of-year grade level appropriate skills. If all students began each grade "on-level," then it might be appropriate to expect all students or a large proportion of students to pass the end-of-grade BSAP test. However, all students do not begin each grade ready to do grade level work. This student entry level is a particular problem at the upper grades.

For example, there were more than 8,000 sixth graders eligible for compensatory and remedial reading programs in 1985. These students had not mastered the material presented in grades 1 through 5. What would have to happen for these students to reach the standard on the BSAP test in the sixth grade? For these students to reach the grade level standard at the end of grade 6, their teachers must help them correct all existing deficiencies from the earlier grades and then help them learn all the grade level material up to grade level standard. This task is difficult at any grade but becomes increasingly difficult as weak students acquire cumulative deficiencies when they move from one grade to another. For these reasons, it seems clearly inappropriate to evaluate BSAP performance against an absolute criterion of a certain percentage of students reaching

grade-level standard. Change in the percentage of students over the years reaching the grade level standard seems to be the more appropriate criterion for assessing the effects of the BSAP program on student learning.

A final issue to be considered while examining the results of the study is the relationship of the BSAP tests to the BSAP instructional objectives. The first major step in the development of the BSAP program was the identification of statewide reading objectives and subskills and the approval of these objectives and subskills as defining grade level appropriate skills. The BSAP tests were constructed to reflect students' performance relative to these skills. Hence, students' test scores should be interpreted in terms of how well students are learning the skills specified as BSAP instructional objectives. Interpreting students' tests scores in terms of how "hard" or "easy" the tests are for the students misses the whole point of the BSAP program and ignores the value of skill-referenced tests. For example, the reader will soon see that a large proportion of first grade students reach the grade level standard in reading. This result indicates that students are learning the skills specified as grade 1 reading objectives. Thus, interpretations and actions based on this finding should focus primarily on the current easyness of first grade objectives and whether the objectives may need to be reconsidered because teachers have become so effective in Teaching them. Action focused on simply making the first grade test "harder" would totally ignore the instructional issues implied by the students' test performance and dismiss the progress made by students as a testing issue not an instructional outcome.

Grade 8 results provide a different example of this important issue of considering the objectives reflected in the test scores. In grade 8 there

is a relatively low proportion of students reaching the grade level standard. The important implication of this finding should focus on the appropriateness of the grade 8 objectives. It is not useful to consider simply making the eighth grade tests "easier" because the tests are simply reflecting the difficulty of the instructional objectives. Concerns about student performance in the upper grades should focus on the nature and appropriateness of the sixth and eighth grade reading objectives and the nature and effectiveness of the instruction students receive. The objectives become broader in scope and more complex as students move up the grades and this changing nature of the reading objectives makes the grade level skills more and more difficult to achieve in successive years.

The reading objectives do not increase in difficulty by an equal amount from one grade to the next; the grade 6 and grade 8 BSAP reading tests are harder than the tests at grades 1, 2 and 3 because the objectives they assess are much more complex and difficult than the objectives assessed at the earlier grades.

RESULTS

The results of this study are organized around the four research questions described in the Introduction. Specific procedures used to answer each question are given before the results are presented and a summary follows the presentation of results for each question.

Overall Student Performance On BSAP Examinations

Question 1. Have students in South Carolina made progress toward achieving the grade level standards in reading and mathematics from 1981 to 1985?

This question is answered by examining the proportion of students in the +R+M groups for 1981 through 1985. In addition to answering Question 1, this general overview of students' performance on BSAP Reading and Mathematics Tests provides an important context for examining students' specific deficiencies in reading. The BSAP performance of students in grades 1, 2, 3, 6 and 8 for 1981 through 1985 is shown in Table 1 on the next page. For each grade, the total number of students tested and the percentage of students reaching standard in both areas (+R+M), in one but not the other area (+R-M or -R+M), and the percentage of students who reached neither standard (-R-M) are reported.

The data in Table 1 indicate that the percentage of students who reached both reading and mathematics standards has increased in all grades across the five years and, conversely, the percentage of students who did not reach either standard has decreased for all grades across the five year

Number (and Percentage) of Students Meeting or Failing to Meet
BSAP Standards in Reading and Mathematics

Grade 1	All Ss	+R+M	-R+M	+R-M	-R-M
1985	49,782	37,754 (75.8%)	3,806 (7.7%)	2,830 (5.7%)	5,392 (10.8%)
1984	49,572	36,304 (73.2%)	3,937 (8.0%)	3,364 (6.8%)	5,967 (12.0%)
1983	46,440	31,207 (67.2%)	4,015 (8.7%)	3,782 (8.1%)	7,436 (16.0%)
1982	46,516	28,076 (60.4%)	3,379 (7.3%)	5,409 (11.6%)	9,652 (20.8%)
1981	47,412	28,245 (59.6%)	4,190 (8.8%)	4,908 (10.4%)	10,069 (21.2%)
Grade 2	**All Ss**	**+R+M**	**-R+M**	**+R-M**	**-R-M**
1985	44,944	33,076 (73.6%)	5,356 (11.9%)	2,501 (5.6%)	4,011 (8.9%)
1984	42,590	29,698 (69.7%)	5,415 (12.7%)	2,469 (5.8%)	5,008 (11.8%)
1983	43,239	27,312 (63.2%)	5,552 (12.8%)	3,124 (7.2%)	7,251 (16.8%)
1982	44,226	24,641 (55.7%)	3,865 (8.7%)	5,660 (12.8%)	10,060 (22.8%)
1981	45,491	24,728 (54.4%)	6,865 (15.1%)	3,273 (7.2%)	10,620 (23.3%)
Grade 3	**All Ss**	**+R+M**	**-R+M**	**+R-M**	**-R-M**
1985	42,219	30,513 (72.3%)	2,952 (7.0%)	4,855 (11.5%)	3,899 (9.2%)
1984	42,840	29,402 (68.6%)	4,447 (10.4%)	3,554 (8.3%)	5,437 (12.7%)
1983	43,957	28,246 (64.3%)	4,465 (10.2%)	4,939 (11.2%)	6,307 (14.4%)
1982	45,207	25,555 (56.5%)	5,062 (11.2%)	5,624 (12.4%)	8,966 (19.8%)
1981	46,952	24,743 (52.7%)	4,065 (8.7%)	6,952 (14.8%)	11,192 (23.8%)
Grade 6	**All Ss**	**+R+M**	**-R+M**	**+R-M**	**-R-M**
1985	44,923	23,647 (52.6%)	3,949 (8.8%)	6,309 (14.0%)	11,018 (24.5%)
1984	46,515	23,015 (49.5%)	3,433 (7.4%)	7,176 (15.4%)	12,891 (27.7%)
1983	49,987	24,093 (48.2%)	4,117 (8.2%)	6,458 (12.9%)	15,319 (30.7%)
1982	48,127	21,827 (45.4%)	2,526 (5.3%)	8,190 (17.0%)	15,584 (32.4%)
1981	46,546	18,764 (40.3%)	3,278 (7.0%)	6,832 (14.7%)	17,672 (38.0%)
Grade 8	**All Ss**	**+R+M**	**-R+M**	**+R-M**	**-R-M**
1985	48,955	24,370 (49.8%)	3,723 (7.6%)	6,998 (14.3%)	13,864 (28.3%)
1984	47,135	21,530 (45.7%)	3,867 (8.2%)	6,807 (14.4%)	14,931 (31.7%)
1983	45,969	16,868 (36.7%)	2,468 (5.4%)	9,032 (19.7%)	17,601 (38.2%)
1982	45,504	16,033 (35.2%)	2,452 (5.4%)	7,744 (17.0%)	19,275 (42.4%)
1981	46,667	17,024 (36.5%)	2,901 (6.2%)	6,795 (14.6%)	19,947 (42.7%)

Code

+R+M: Student Passing Both Tests
-R+M: Students Failing Reading, Passing Mathematics
+R-M: Students Passing Reading, Failing Mathematics
-R-M: Students Failing Both Tests

INCREASED AND DECREASED PERCENTAGE OF STUDENTS
IN THE +R+M AND -R-M GROUPS FROM 1981 TO 1985

TABLE 2

Grade	Increased Percentage of Students in the +R+M Group	Decreased Percentage of Students in the -R-M Group	Total Shift in Student Population
1	+16.2	-10.4	26.6
2	+19.2	-14.4	33.6
3	+19.6	-14.6	34.2
6	+12.3	-13.6	25.9
8	+13.3	-14.4	27.7
Average	+16.1	-13.5	29.6

period. These increases in +R+M percentages and decreases in -R-M percentages are summarized for each grade in Table 2 on page 16. The increase in the percentage of students in the +R+M group is greatest for grade 3 with an increase of +19.6%, and least for grade 6 with a +12.3% increase. The average increase is +16.1% for all grades. The decreased percentage of students in the -R-M category is consistent across the five grades with an average decrease of -13.5%. The total shift in student population is determined by adding the increased percentage of students in the +R+M and the decreased percentage of students from out of the -R-M group. This total population shift varies from 25.9% in grade 6 to 34.2% in grade 3, with an average shift of 29.6% for all grades.

Discussion and Interpretation

The data in Tables 1 and 2 are quite clear. Across all grades there is a dramatic increase in the proportion of students meeting both the reading and mathematic standards and a dramatic decrease in the proportion of students who reach neither standard. Many factors may contribute to positive shifts of 26% to 34% for an entire school population, but the single most likely explanation for these gains is an increased instructional effort directed toward the statewide basic skills objectives in reading and mathematics. Increased program implementation by teachers, schools, and school districts since the BSAP program began in 1981 is almost surely the major force behind the increased percentage of students meeting both standards and the decreased percentage of students meeting neither standard.

The cumulative effect of instruction aimed at the BSAP Reading and Mathematics Objectives also is likely to be reflected in the performance of students in grades 1, 2 and 3, but would not yet be apparent for students in grades 6 and 8. The first "BSAP Class" were first graders in 1981 and

had only reached the fifth grade by 1985. Therefore, no students in grades 6 and 8 had been exposed to BSAP oriented instruction throughout their school years. In contrast, the third grade students had been exposed to BSAP oriented instruction from the time they began school. The increased percentage of students meeting both standards is greatest for grade 3 (19.6%) and then grade 2 (19.2%) and then grade 1 (16.2%). This ordering may reflect the cumulative effect of BSAP oriented instruction. The relatively smaller amount of progress in grades 6 and 8 may also reflect the previously mentioned issue of increased complexity of learning objectives and decreased instructional emphasis on basic skills at these grades.

A More Detailed Examination Of Students' Overall Performance

Question 2. *Do the students who achieve both the reading and mathematics standards in 1985 reach the reading standard by as much as the corresponding students in 1981? And, are the students who reach neither the reading nor mathematics standards in 1985 any closer to reaching the reading standard than the corresponding students in 1981?*

A central finding of this study described in the preceding section is that from 1981 to 1985 at all grade levels there are more students doing well, as evidenced by an increased percentage of students in the +R+M group, and fewer students doing poorly, as evidenced by the decreased percentage of students in the -R-M group. A more detailed examination of students' improved achievement on BSAP reading examinations can be obtained by determining whether the students who passed both reading and mathematics

standards in 1985 are passing the reading standards by as much as their 1981 counterparts passed the standard. Likewise, it is informative to determine whether the students who did not reach either reading or mathematics standards in 1985 are any closer to reaching the reading standard than their 1981 counterparts.

The standard error of measurement (S.E.M.) is used to make these determinations. The S.E.M. is a number of points on a scale that is used to construct a confidence interval around a test score. An interval of plus and minus one S.E.M. around a score gives a 68% confidence interval; i.e., there is a 68% chance that a person's true score falls in the interval. Plus and minus two S.E.M.'s gives a 95% confidence interval; i.e., there is a 95% chance that a person's true score falls in the interval.

The use of S.E.M's to answer Question 2 can be illustrated by considering six hypothetical students. The performance of these hypothetical students is shown in Figure 3 on the next page. Student A passes the standard by greater than two S.E.M.'s and student B passes the standard by greater than one S.E.M. but less than 2 S.E.M.'s. Both of these students have passed the standard, but student A has passed it at a higher level than student B. Student C also passes the standard but does so by less than one S.E.M. Student C has passed the standard but has done less well than the other two students.

On the other side of the standard consider students X, Y and Z, all of whom fail to make the standard. Student X is less than one S.E.M. below standard, student Y is greater than one but less than two S.E.M.'s below standard, and student Z is greater than two S.E.M.'s below standard. All three students have failed to make the standard. However, student X did

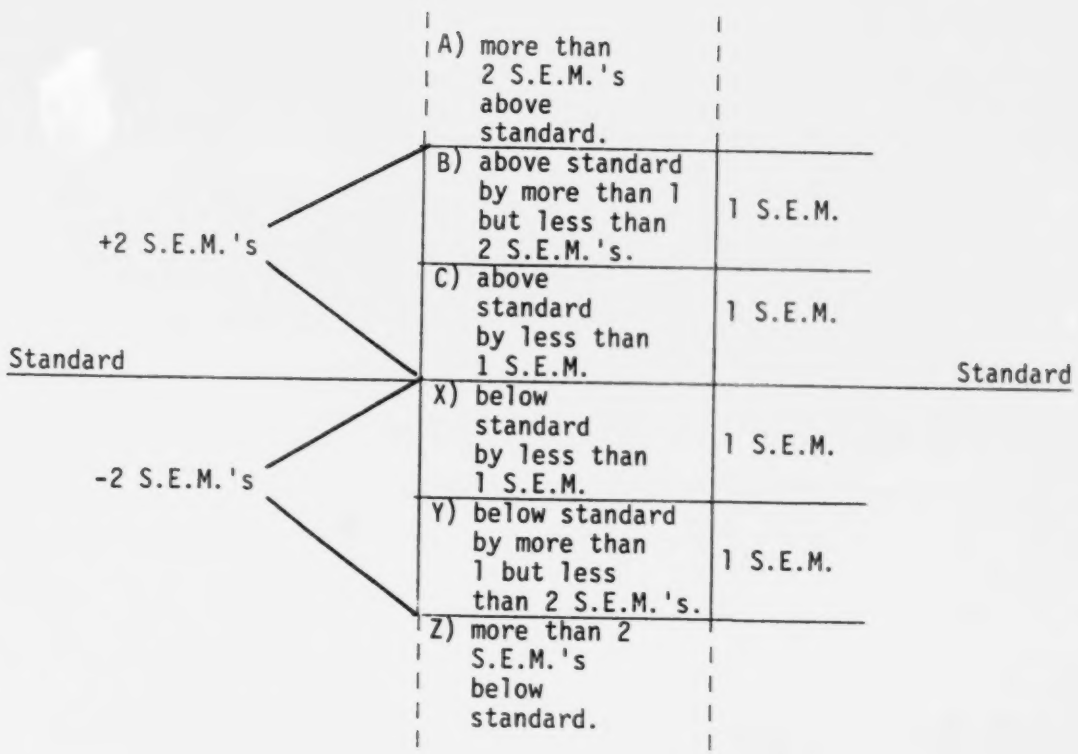

A) more than 2 S.E.M.'s above standard.
B) above standard by more than 1 but less than 2 S.E.M.'s.
1 S.E.M.
+2 S.E.M.'s
C) above standard by less than 1 S.E.M.
1 S.E.M.
Standard
Standard
X) below standard by less than 1 S.E.M.
1 S.E.M.
-2 S.E.M.'s
Y) below standard by more than 1 but less than 2 S.E.M.'s.
1 S.E.M.
Z) more than 2 S.E.M.'s below standard.

Figure 3: Illustration of Classification of Students by S.E.M.'s

not miss the standard by as much as student Y, and student Y was closer to making the standard than student Z.

The approach illustrated with these six hypothetical students is used to determine if the +R+M students in 1985 are passing the reading standards by as much as the 1981 +R+M students and also to determine if the -R-M students in 1985 are any closer to making the reading standard than the 1981 -R-M students. The distribution in percentages of students' scores on the BSAP reading tests according to their distance above or below the standard for each test is given in the tables beginning on the page 23. Each page displays the information for one grade, with the score distribution for the 1985, 1984, 1983, 1982 and 1981 tests presented from the top to the bottom of the page. The table for each grade presents the percent of students in five groups, the groups being composed of all students combined, +R+M, +R-M, -R+M, and -R-M students by year from 1981 through 1985. The distances of the students' scores from the standard are expressed in the standard errors of measurement (S.E.M.) arranged in the following order from the top to the bottom of the table for each year:

a) greater than two S.E.M. above the standard, (abbreviated: GT.2 above);

b) greater than one and less than two S.E.M. above the standard, (abbreviated: GT.1 & LT.2);

c) scoring at the standard or above it by less than one S.E.M., abbreviated: at or LT.1);

x) below the standard by less than one S.E.M., (abbreviated: below by LT.1);

y) greater than one but less than two S.E.M. below the

standard, (abbreviated: GT.1 but LT.2);

z) greater than two S.E.M. below the standard, (abbreviated: GT.2 below).

The percentage of all students who reached or did not reach the standard is listed in the middle row for each year. The total number and percentages of students in each group are listed in the last two rows for each year.

To illustrate the use of these tables, consider grade 1 (Table 3) for the year 1981 at the bottom of the page. The total number of first grade students (All Ss column) taking the test is 47,412 as listed on the second to the last line on the page. There were 28,245 students or 59.6% of all first graders in the +R+M group. Of these 28,245 students, 81.9% passed the reading standard by greater than two (GT.2 above) S.E.M.'s; 7.6% of the students passed by greater than one but less than two (GT.1 & LT.2) S.E.M.'s; and 10.5% were at or above the standard by less than one (LT.1) S.E.M.

The corresponding group of students five years later is found at the top of the page for 1985. In 1985 there were 37,754 students or 75.8% of all first graders in the +R+M group. This is an increase of 16.2% over 1981 in the percentage of students in the +R+M group. Of these 37,754 first grade students in the 1985 +R+M group, 85.9% passed the reading standard by greater than two S.E.M.'s. This 85.9% is an increase of 4.0% over 1981 in the percentage of students in the +R+M group who passed the standard at a very high level. Thus, not only is there an increase of 16.2% in the percentage of students in the +R+M group from 1981 to 1985, but also within this group there is an increase of 4.0% in the percentage of +R+M youngsters who exceed the standard by more than two S.E.M.'s.

Distribution of Students by Group Membership
on BSAP Reading

Grade 1

1985

S.E.M. above or below	All Ss	+R-M	-R+M	+R-M	-R-M
a)GT.2 above	67.5	85.9		42.1	
b)GT.1 & LT.2	5.7	6.1		18.3	
c)at or LT.1	8.3	8.0		39.6	
Pass/Fail	81.5/18.5				
x)below by LT.1	4.9		38.4		18.2
y)GT.1 but LT.2	4.4		28.5		20.6
z)GT.2 below	9.2		33.1		61.2
TOTAL	49782	37754	3806	2830	5392
Percent of total	100.0	75.8	7.7	5.7	10.8

1984

S.E.M. above or below	All Ss	+R-M	-R+M	+R-M	-R-M
a)GT.2 above	65.4	85.1		46.1	
b)GT.1 & LT.2	6.0	6.6		17.3	
c)at or LT.1	8.6	8.3		36.6	
Pass/Fail	80.0/20.0				
x)below by LT.1	5.1		37.5		17.7
y)GT.1 but LT.2	4.9		29.7		21.2
z)GT.2 below	10.0		32.8		61.1
TOTAL	49572	36304	3937	3364	5967
Percent of total	100.0	73.2	8.0	6.8	12.0

1983

S.E.M. above or below	All Ss	+R-M	-R+M	+R-M	-R-M
a)GT.2 above	60.9	84.9		47.9	
b)GT.1 & LT.2	5.8	6.5		17.6	
c)at or LT.1	8.6	8.6		34.5	
Pass/Fail	75.3/24.7				
x)below by LT.1	5.7		35.5		16.3
y)GT.1 but LT.2	5.6		28.1		19.6
z)GT.2 below	13.4		36.4		64.1
TOTAL	46440	31207	4015	3782	7436
Percent of total	100.0	67.2	8.7	8.1	16.0

1982

S.E.M. above or below	All Ss	+R+M	-R+M	+R-M	-R-M
a)GT.2 above	57.6	85.8		50.0	
b)GT.1 & LT.2	5.5	6.1		15.8	
c)at or LT.1	8.9	8.1		34.2	
Pass/Fail	72.0/28.0				
x)below by LT.1	6.4		37.6		17.7
y)GT.1 but LT.2	6.0		27.8		19.4
z)GT.2 below	15.6		34.6		62.9
TOTAL	46561	28076	3379	5409	9652
Percent of total	100.0	60.4	7.3	11.6	20.8

1981

S.E.M. above or below	All Ss	-R-M	-R+M	+R-M	-R-M
a)GT.2 above	53.2	81.9		42.3	
b)GT.1 & LT.2	6.3	7.6		17.4	
c)at or LT.1	10.4	10.5		40.3	
Pass/Fail	69.9/30.1				
x)below by LT.1	7.2		38.9		17.7
y)GT.1 but LT.2	6.9		28.3		20.5
z)GT.2 below	16.0		32.8		61.8
TOTAL	47412	28245	4190	4908	10069
Percent of total	100.0	59.6	8.8	10.4	21.2

Distribution of Students by Group Membership
on BSAP Reading

Grade 2

1985

S.E.M. above or below	All Ss	+R+M	-R+M	+R-M	-R-M
a)GT.2 above	67.8	87.7		57.4	
b)GT.1 & LT.2	6.7	7.3		23.6	
c)at or LT.1	4.7	5.0		19.0	
Pass/Fail	79.2/20.8				
x)below by LT.1	1.9		11.9		5.4
y)GT.1 but LT.2	3.4		20.4		11.4
z)GT.2 below	15.5		67.7		83.2
TOTAL	44944	33076	5356	2501	4011
Percent of total	100.0	73.6	11.9	5.6	8.9

1984

S.E.M. above or below	All Ss	+R+M	-R+M	-R-M	-R-M
a)GT.2 above	55.9	77.1		37.8	
b)GT.1 & LT.2	9.9	12.1		25.2	
c)at or LT.1	9.7	10.8		37.0	
Pass/Fail	75.5/24.5				
x)below by LT.1	4.6		25.5		11.7
y)GT.1 but LT.2	3.9		19.8		11.8
z)GT.2 below	16.0		54.7		76.5
TOTAL	42590	29698	5415	2469	5008
Percent of total	100.0	69.7	12.7	5.8	11.8

1983

S.E.M. above or below	All Ss	+R+M	-R+M	+R-M	-R-M
a)GT.2 above	50.2	75.4		34.8	
b)GT.1 & LT.2	10.1	13.0		26.8	
c)at or LT.1	10.1	11.6		38.4	
Pass/Fail	70.4/29.6				
x)below by LT.1	5.0		25.0		10.9
y)GT.1 but LT.2	4.3		19.3		10.8
z)GT.2 below	20.3		55.7		78.3
TOTAL	43239	27319	5552	3124	7251
Percent of total	100.0	63.2	12.8	7.2	16.8

1982

S.E.M. above or below	All Ss	+R+M	-R+M	-R-M	-R-M
a)GT.2 above	51.9	81.8		48.9	
b)GT.1 & LT.2	8.4	10.0		22.4	
c)at or LT.1	8.2	8.2		28.7	
Pass/Fail	68.5/31.5				
x)below by LT.1	4.3		23.1		10.1
y)GT.1 but LT.2	3.8		18.1		9.7
z)GT.2 below	23.4		58.8		80.2
TOTAL	44226	24641	3865	5660	10060
Percent of total	100.0	55.7	8.7	12.8	22.8

1981

S.E.M. above or below	All Ss	+R+M	-R+M	+R-M	-R-M
a)GT.2 above	44.5	76.5		40.9	
b)GT.1 & LT.2	8.2	11.8		24.0	
c)at or LT.1	8.9	11.7		35.1	
Pass/Fail	61.6/38.4				
x)below by LT.1	4.7		19.5		7.8
y)GT.1 but LT.2	4.3		16.1		7.9
z)GT.2 below	29.4		64.4		84.3
TOTAL	45491	24728	6865	3278	10620
Percent of total	100.0	54.4	15.1	7.2	23.3

Distribution of Students by Group Membership
on BSAP Reading

Grade 3

1985

S.E.M. above or below	All Ss	+R+M	-R+M	+R-M	-R-M
a)GT.2 above	61.1	80.5		44.5	
b)GT.1 & LT.2	10.6	10.8		24.3	
c)at or LT.1	9.9	8.7		31.2	
Pass/Fail	83.8/16.2				
x)below by LT.1	4.1		31.9		20.0
y)GT.1 but LT.2	3.0		21.1		16.0
z)GT.2 below	9.2		46.9		64.0
TOTAL	42219	30513	2952	4855	3899
Percent of total	100.0	72.3	7.0	11.5	9.2

1984

S.E.M. above or below	All Ss	+R+M	-R+M	+R-M	-R-M
a)GT.2 above	51.9	72.0		30.6	
b)GT.1 & LT.2	12.4	15.0		24.8	
c)at or LT.1	12.6	13.0		44.6	
Pass/Fail	76.9/23.1				
x)below by LT.1	6.0		35.3		18.8
y)GT.1 but LT.2	4.6		23.1		17.1
z)GT.2 below	12.5		41.6		64.1
TOTAL	42840	29402	4447	3552	5437
Percent of total	100.0	68.6	10.4	8.3	12.7

1983

S.E.M. above or below	All Ss	+R+M	-R+M	+R-M	-R-M
a)GT.2 above	51.4	73.6		36.9	
b)GT.1 & LT.2	11.6	13.7		24.8	
c)at or LT.1	12.5	12.7		38.3	
Pass/Fail	75.5/24.5				
x)below by LT.1	5.6		30.3		17.4
y)GT.1 but LT.2	4.2		21.7		14.3
z)GT.2 below	14.7		48.0		68.3
TOTAL	41957	28246	4465	4939	6307
Percent of total	100.0	64.3	10.2	11.2	14.4

1982

S.E.M. above or below	All Ss	+R+M	-R+M	+R-M	-R-M
a)GT.2 above	39.6	64.1		26.5	
b)GT.1 & LT.2	14.4	19.2		28.9	
c)at or LT.1	15.0	16.7		44.6	
Pass/Fail	69.0/31.0				
x)below by LT.1	7.2		33.6		17.6
y)GT.1 but LT.2	5.2		21.0		14.4
z)GT.2 below	18.6		45.4		68.0
TOTAL	45207	25555	5062	5624	8966
Percent of total	100.0	56.5	11.2	12.4	19.8

1981

S.E.M. above or below	All Ss	+R+M	-R+M	+R-M	-R-M
a)GT.2 above	45.9	75.8		39.8	
b)GT.1 & LT.2	10.4	12.8		25.0	
c)at or LT.1	11.2	11.4		35.2	
Pass/Fail	67.5/32.5				
x)below by LT.1	5.6		28.1		13.4
y)GT.1 but LT.2	4.7		19.6		12.4
z)GT.2 below	22.2		52.3		74.2
TOTAL	46952	24743	4065	6952	11192
Percent of total	100.0	52.7	8.7	14.8	23.8

Distribution of Students by Group Membership
on BSAP Reading

Grade 6

1985

S.E.M. above or below	All Ss	+R+M	-R+M	+R-M	-R-M
a)GT.2 above	42.9	73.1		31.8	
b)GT.1 & LT.2	10.3	18.3		23.8	
c)at or LT.1	13.4	13.6		44.4	
Pass/Fail	66.7/33.3				
x)below by LT.1	8.3		40.0		19.6
y)GT.1 but LT.2	7.0		25.7		19.5
z)GT.2 below	18.0		34.3		60.9
TOTAL	44923	23647	3949	6309	11018
Percent of total	100.0	52.6	8.8	14.0	24.5

1984

S.E.M. above or below	All Ss	+R+M	-R+M	+R-M	-R-M
a)GT.2 above	41.3	73.5		32.1	
b)GT.1 & LT.2	10.3	13.0		24.8	
c)at or LT.1	13.3	13.5		43.1	
Pass/Fail	64.9/35.1				
x)below by LT.1	7.9		39.5		17.9
y)GT.1 but LT.2	6.9		25.4		18.1
z)GT.2 below	20.3		35.1		64.0
TOTAL	46515	23015	3433	7176	12891
Percent of total	100.0	49.5	7.4	15.4	27.7

1983

S.E.M. above or below	All Ss	+R+M	-R+M	+R-M	-R-M
a)GT.2 above	30.0	57.7		16.5	
b)GT.1 & LT.2	16.7	26.0		32.6	
c)at or LT.1	14.4	16.3		50.9	
Pass/Fail	61.1/38.9				
x)below by LT.1	8.8		40.8		17.6
y)GT.1 but LT.2	11.0		35.1		26.6
z)GT.2 below	19.1		24.1		55.8
TOTAL	49987	24093	4117	6458	15319
Percent of total	100.0	48.2	8.2	12.9	30.7

1982

S.E.M. above or below	All Ss	+R+M	-R+M	+R-M	-R-M
a)GT.2 above	40.6	76.6		34.8	
b)GT.1 & LT.2	9.3	11.9		22.7	
c)at or LT.1	12.5	11.5		42.5	
Pass/Fail	62.4/37.6				
x)below by LT.1	7.4		38.9		16.4
y)GT.1 but LT.2	6.7		24.7		16.8
z)GT.2 below	23.5		36.4		66.8
TOTAL	48127	21827	2526	8190	15584
Percent of total	100.0	45.4	5.3	17.0	32.4

1981

S.E.M. above or below	All Ss	+R+M	-R+M	+R-M	-R-M
a)GT.2 above	33.7	72.0		31.7	
b)GT.1 & LT.2	8.6	13.3		22.4	
c)at or LT.1	12.7	14.7		45.9	
Pass/Fail	55.0/45.0				
x)below by LT.1	7.9		37.4		13.8
y)GT.1 but LT.2	7.6		25.2		15.3
z)GT.2 below	29.5		37.4		70.9
TOTAL	46546	18764	3278	6832	17672
Percent of total	100.0	40.3	7.0	14.7	38.0

Distribution of Students by Group Membership
on BSAP Reading

Grade 8

1985

S.E.M. above or below	All Ss	+R+M	-R+M	+R-M	-R-M
a)GT.2 above	38.3	69.0		27.6	
b)GT.1 & LT.2	11.9	16.6		25.3	
c)at or LT.1	13.9	14.4		47.1	
Pass/Fail	64.1/35.9				
x)below by LT.1	7.8		38.8		17.1
y)GT.1 but LT.2	6.6		25.0		16.8
z)GT.2 below	21.5		36.2		66.1
TOTAL	48955	24370	3723	6998	13864
Percent of total	100.0	49.8	7.6	14.3	28.3

1984

S.E.M. above or below	All Ss	+R+M	-R+M	+R-M	-R-M
a)GT.2 above	31.4	61.5		22.9	
b)GT.1 & LT.2	13.0	20.2		25.9	
c)at or LT.1	15.7	18.3		51.2	
Pass/Fail	60.1/39.9				
x)below by LT.1	8.6		37.9		17.3
y)GT.1 but LT.2	7.2		24.9		16.4
z)GT.2 below	24.1		37.2		66.3
TOTAL	47135	21530	1867	6807	14931
Percent of total	100.0	45.7	8.2	14.4	31.7

1983

S.E.M. above or below	All Ss	+R+M	-R+M	+R-M	-R-M
a)GT.2 above	22.0	51.9		14.8	
b)GT.1 & LT.2	18.4	30.8		36.2	
c)at or LT.1	15.9	17.3		49.0	
Pass/Fail	56.3/43.7				
x)below by LT.1	8.6		40.6		16.7
y)GT.1 but LT.2	11.0		33.8		24.1
z)GT.2 below	24.1		25.6		59.2
TOTAL	45969	16868	2468	9032	17601
Percent of total	100.0	36.7	5.4	19.7	38.2

1982

S.E.M. above or below	All Ss	+R+M	-R+M	+R-M	-R-M
a)GT.2 above	24.7	60.2		20.6	
b)GT.1 & LT.2	15.3	26.0		35.9	
c)at or LT.1	12.3	13.8		43.5	
Pass/Fail	52.3/47.7				
x)below by LT.1	7.7		37.5		17.5
y)GT.1 but LT.2	10.2		31.4		20.9
z)GT.2 below	29.5		31.1		65.6
TOTAL	45504	16033	2452	7744	19275
Percent of total	100.0	35.2	5.4	17.0	42.4

1981

S.E.M. above or below	All Ss	+R+M	-R+M	+R-M	-R-M
a)GT.2 above	28.8	68.4		26.3	
b)GT.1 & LT.2	9.7	16.3		25.8	
c)at or LT.1	12.5	15.3		47.9	
Pass/Fail	51.0/49.0				
x)below by LT.1	7.5		35.6		12.3
y)GT.1 but LT.2	7.1		23.7		13.2
z)GT.2 below	34.4		40.7		74.5
TOTAL	46667	17024	2901	6795	19947
Percent of total	100.0	36.5	6.2	14.6	42.7

A similar type of analysis can be conducted for the -R-M groups for 1981 and 1985. Again considering first grade students, Table 3 shows that in 1981 there were 10,069 or 21.2% of all first graders in the -R-M group. In 1985, there were 5,392 or 10.8% of all first graders in the same group. Thus, there has been a decrease of 10.4% in the percentage of all students in the -R-M group. In 1981, 61.8% of the students in the -R-M group failed the standard by more than two S.E.M and in 1985, 61.2% of students in the -R-M group failed the standard by more than two S.E.M's. Thus, not only is there a decrease of 10.4% in the percentage of students in the -R-M group, but also within that group there is a decrease of 0.6% in the percentage of students failing the standard by greater than two S.E.M.'s.

A similar analysis can be performed for students in grades 2, 3, 6 and 8 based on the information in the appropriate tables. A summary of these analyses is given in Table 8 on the next page. For each grade, this table shows the change in the percentage of students in the -R-M group who fell below the standard by greater than two S.E.M.'s.

The information in Table 8 indicates that there is an increase in the percentage of +R+M students passing the reading standard by greater than two S.E.M.'s in grades 1, 2, 3, 6 and 8 from 1981 to 1985. The average increase across the grades is 4.3%. Also, in the -R-M group, there is a decrease in the percentage of students failing the reading standard by greater than two S.E.M.'s in all grades from 1981 to 1985. The average decrease across the grades is 6.0%.

1981-1985 Changes in the Percentage of Students
Passing or Failing the Reading Standard
by Greater than two S.E.M.'s

TABLE 8

	1981-1985 Change in the Percentage of +R+M Students Meeting the Standard by Greater Than 2 S.E.M.'s			1981-1985 Change in the Percentage of -R-M Students Failing the Standard by Greater Than 2 S.E.M.'s		
Grade	1981	1985	Change	1981	1985	Change
1	81.9	85.9	4.0	61.8	61.2	- 0.6
2	76.5	87.7	11.2	84.3	83.2	- 1.1
3	75.8	80.5	4.7	74.2	64.0	-10.2
6	72.0	73.1	1.1	70.9	60.9	-10.0
8	68.4	69.0	0.6	74.5	66.1	- 8.4
Average	74.9	79.2	+4.3	73.1	67.1	- 6.0

Discussion and Interpretation

In reflecting on the preceding analysis, it is important to remember earlier information (Table 2) which indicates that in all grades there is an increase in the average percentage of students in the +R+M group (16.1%) and a decrease in the average percentage of students in the -R-M group (-13.5%) from 1981 to 1985. This information considered along with the information from Table 8 means that (1) there are more students in 1985 meeting the grade level standards than there were in 1981 (16.1%), and that (2) these 1985 students are not just barely meeting the standard but they are meeting the standard by a greater amount than the 1981 +R+M students. This description applies to all grades. On the other side of the standard, the data indicate that (1) there are fewer students in 1985 who do not reach both standards than there were in 1981 (-13.5% average), and that (2) the students who reached neither standard in 1985 are closer to the reading standard than the 1981 -R-M students. These findings apply to all grades.

Analyzing Students' Deficiencies In Reading

Question 3. Are there specific areas of reading in which students across the state have deficiencies? If so, what are the deficiency areas?

Three levels of analysis were conducted in an effort to identify students' deficiencies in BSAP Reading skills. These analyses examined students' performance at the objective, subskill and item levels. The item analysis provides information that is likely to be more useful to teachers

than the other two analyses; hence, a major portion of this report is devoted to describing students' deficiencies at the item level.

Students' Performance on the Reading Objectives

The six BSAP Reading objectives are:

DW: Decoding and Word Meaning

DE: Details

MI: Main Idea

RE: Reference Usage

IN: Inference

AL: Analysis of Literature

Student performance on the BSAP Reading objectives is examined by calculating an overall difficulty for each objective. This objective difficulty level is calculated in two steps. In the first step, the proportion of students correctly answering each item used to assess a particular objective is calculated. This proportion is called the item difficulty and is expressed as the proportion or percentage of students who correctly answer the item. An item difficulty of .90 means that 90% of the students answered the item correctly, while an item difficulty of .40 means that 40% of the students answered the item correctly. An item with a difficulty of .90 is clearly much easier than one with a difficulty of .40.

The second step in calculating an objective difficulty entails calculating the average of the item difficulties. For example, an objective which is assessed by two items with item difficulties of .90 and .40 would

have an objective difficulty of .65. There are actually six items per objective for each of the five years, or 30 items assessing each objective.

The objective difficulties are calculated separately for the +R and -R students. The difference between the objective difficulty for the +R group and the objective difficulty for the -R group indicates how well the objective differentiates or discriminates between the two groups.

The results of these analysis are shown in Figures 4 through 8 on the pages that follow for grades 1, 2, 3, 6 and 8, respectively. The results for each grade will be briefly described, followed by a description of some general findings seen across all grades.

Grade 1

The analysis of Grade 1, Figure 4, shows that for the +R students the objective difficulties range from 96.9% for Decoding and Word Meaning (DW) to 76.8% for Analysis of Literature (AL). The objectives are fairly well bunched together for the +R students indicating homogeneity of objective difficulty.

For the -R students the objective difficulties are much lower than for the +R students. The best performance of the -R students is 71.8% on DW, which is lower than the worst performance of the +R students (76.8% on AL). Thus all the objectives are sharply differentiating the -R from the +R students. The performance of the -R students on AL (32.1%) and DE (33.3%) is very low since at grades 1, 2, and 3 each question has only 3 options. With 3-option multiple choice items, an objective difficulty of 33.3% can

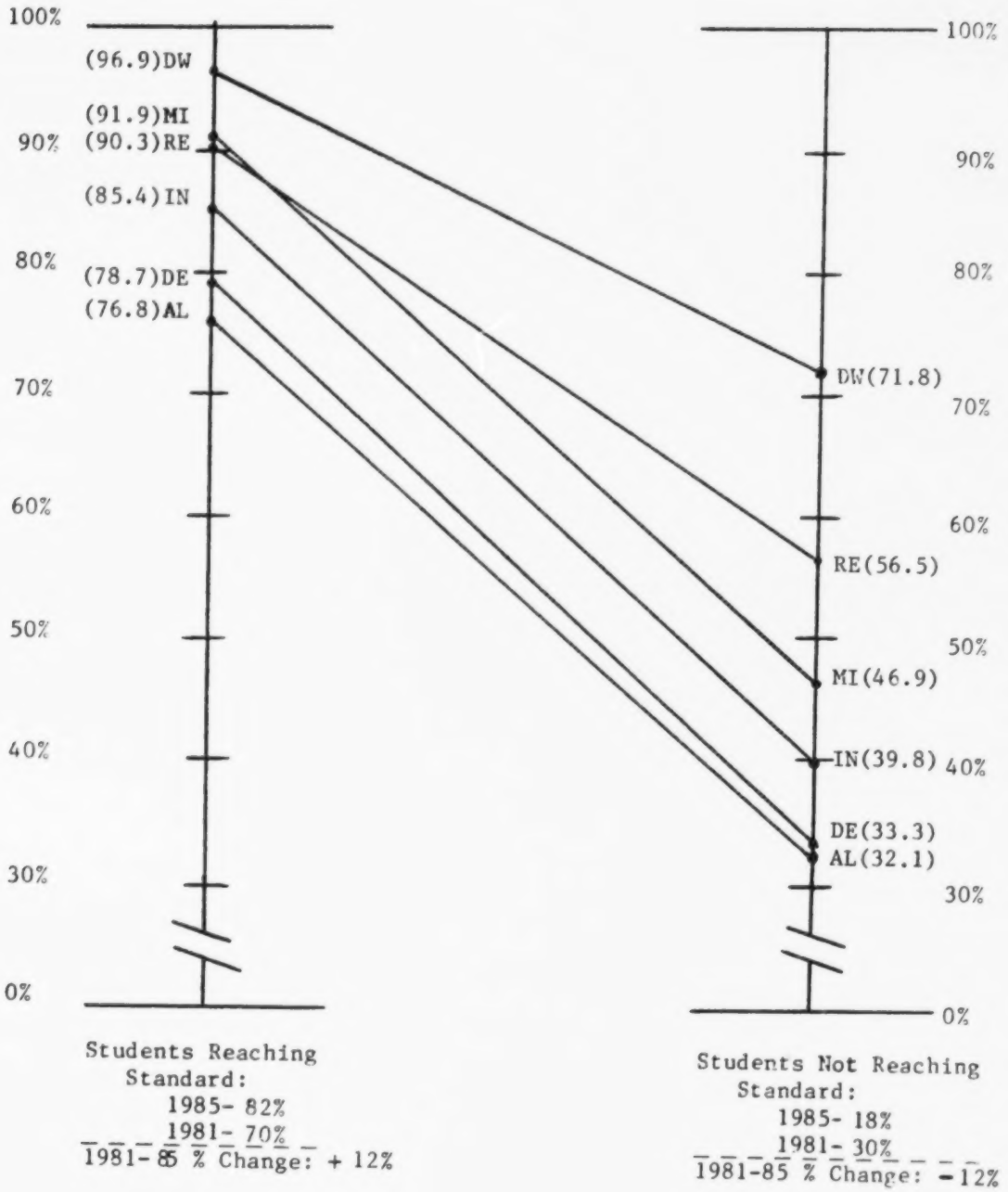


Figure 4. : Average percentage of Grade 1 students reaching and not reaching standard who correctly answer items in the six BSAP objectives

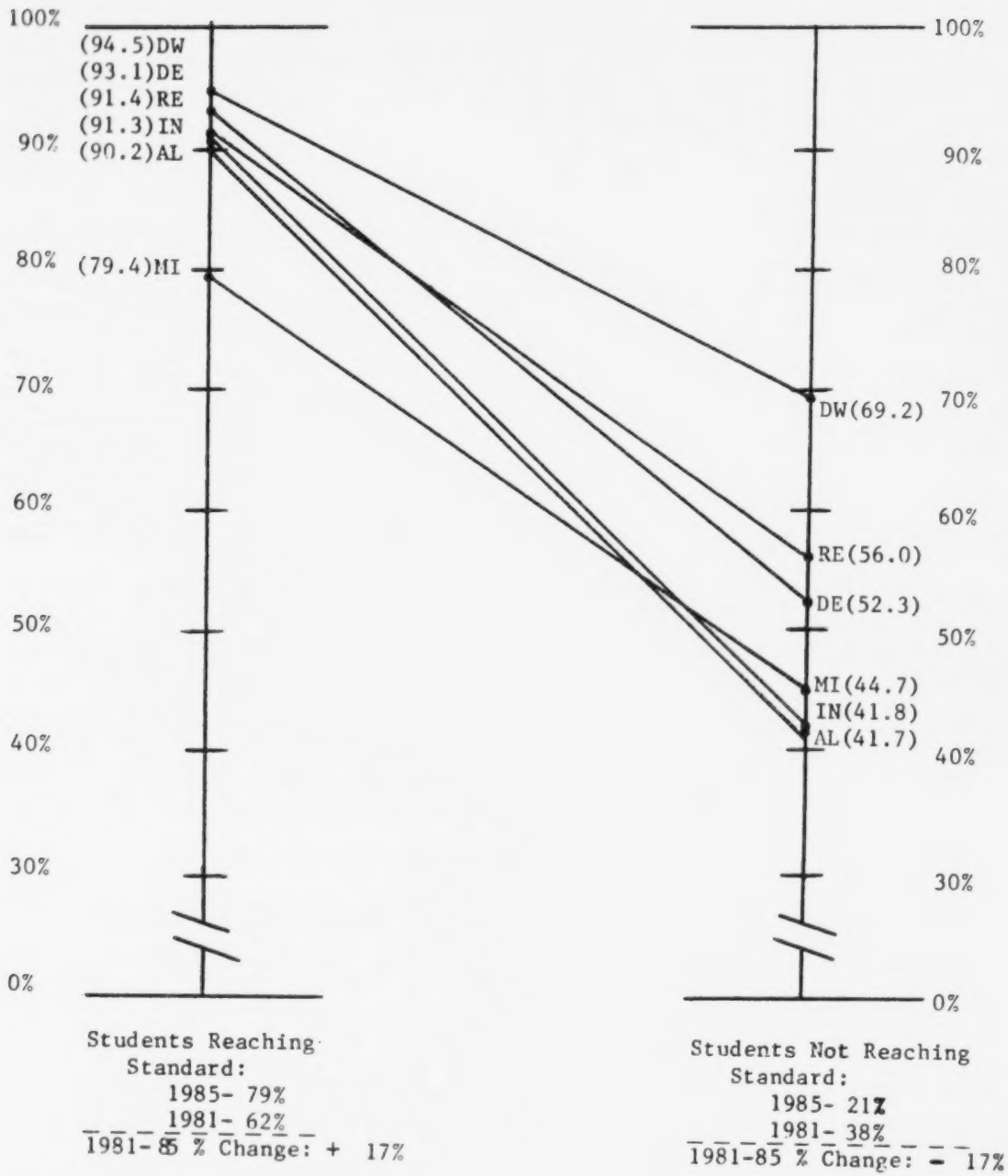


Figure 5. : Average percentage of Grade 2 students reaching and not reaching standard who correctly answer items in the six BSAP objectives

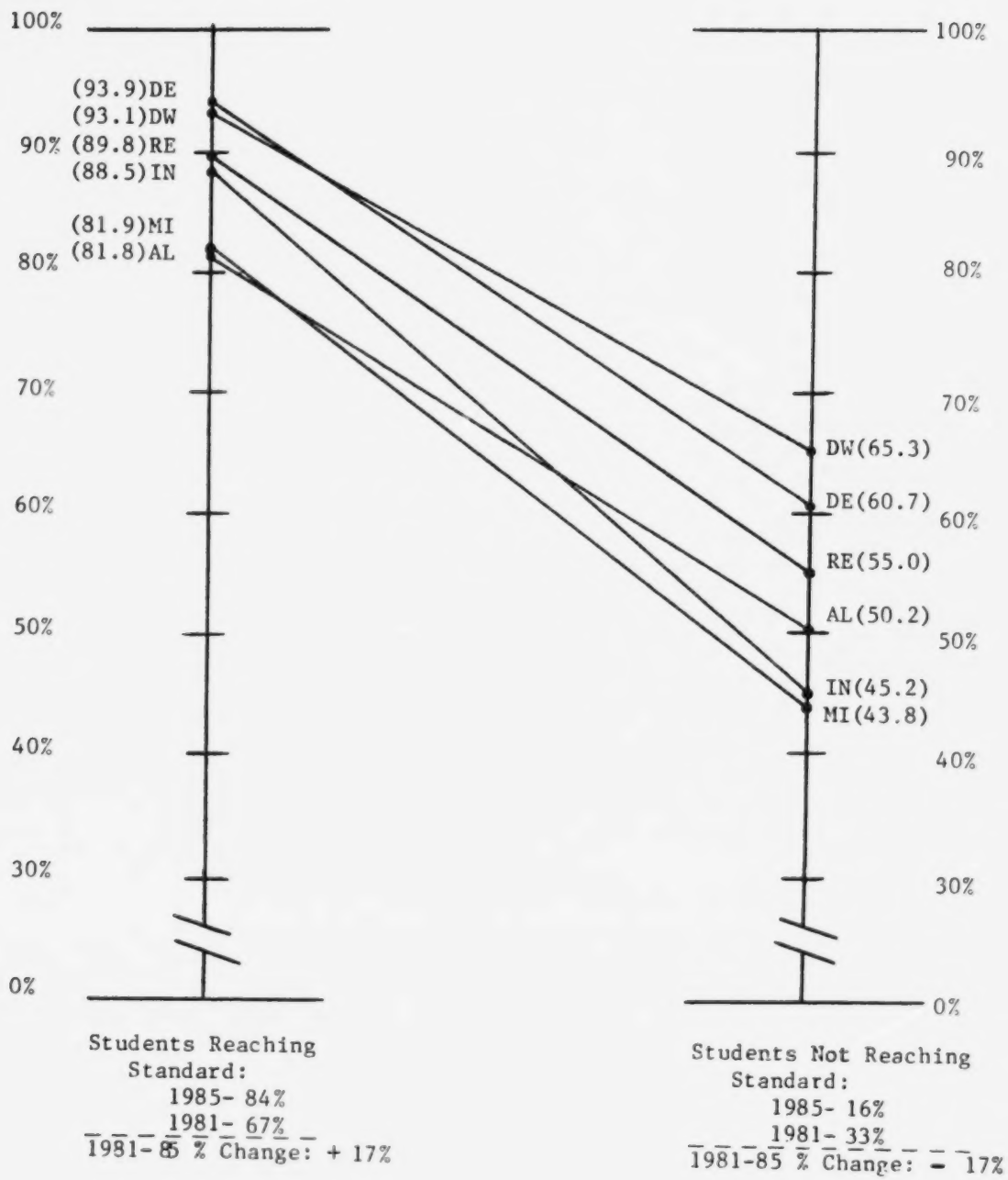


Figure 6. : Average percentage of Grade 3 students reaching and not reaching standard who correctly answer items in the six BSAP objectives

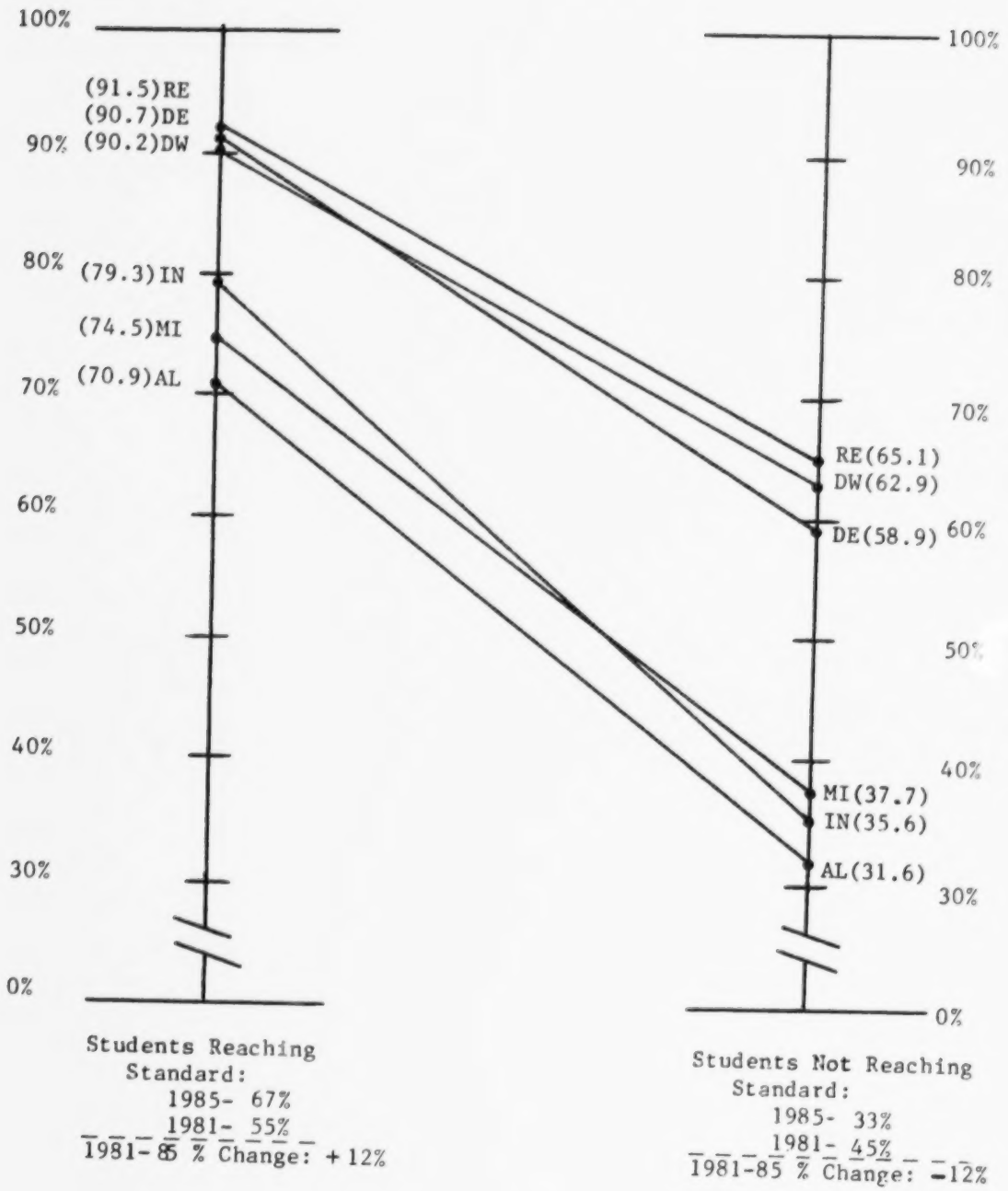


Figure 7. : Average percentage of Grade 6 students reaching and not reaching standard who correctly answer items in the six BSAP objectives

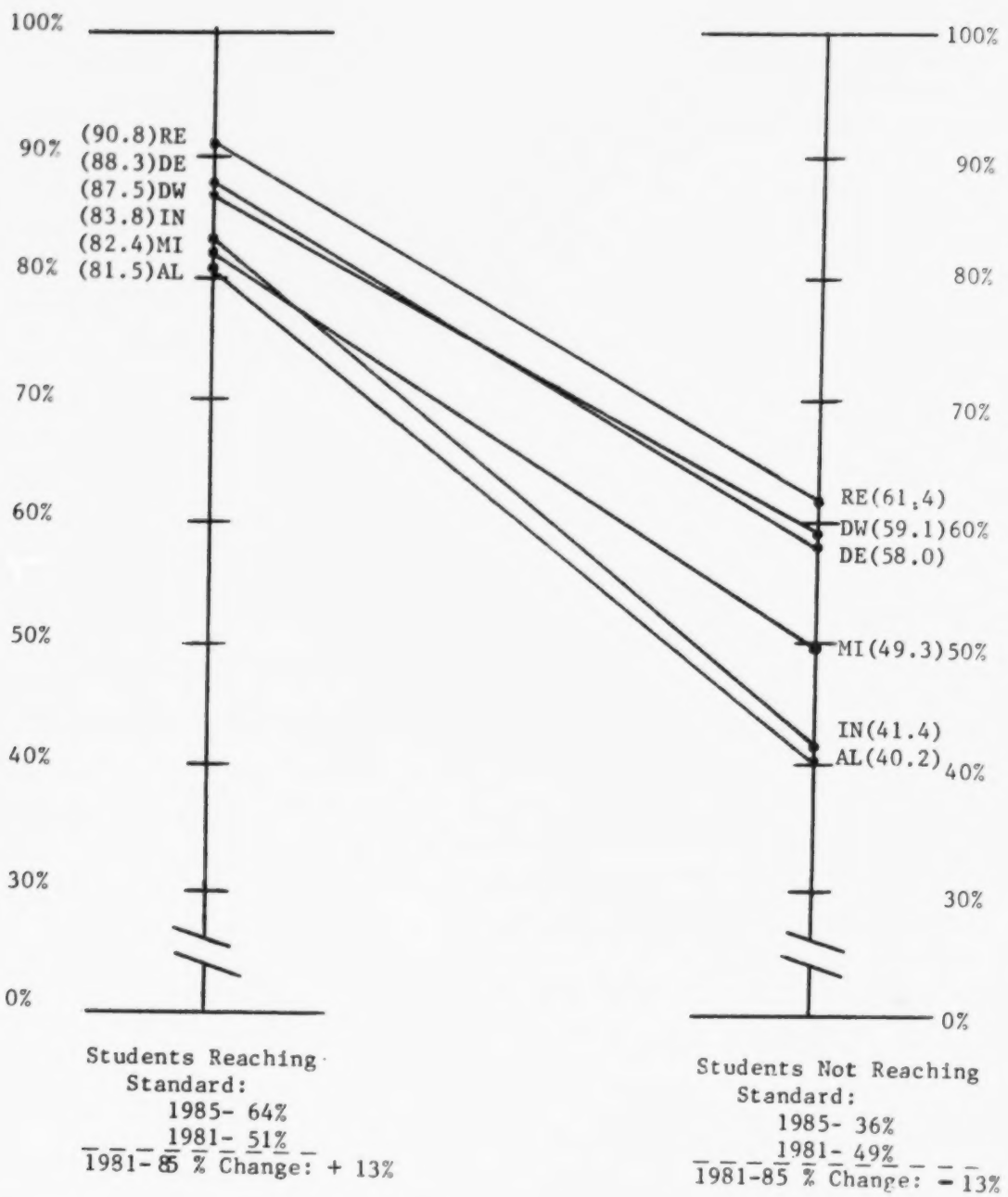


Figure 8. : Average percentage of Grade 8 students reaching and not reaching standard who correctly answer items in the six BSAP objectives

be achieved by randomly guessing at the answers. The performance of the -R students on DE and AL is at or below this chance level which suggests that they have very serious deficiencies in these objectives and/or they do not understand what they are to do in answering the questions. The ordering or relative difficulty of the objectives for the +R and -R students is very consistent.

Grade 2

At grade 2, five of the objectives (See Figure 5) have very similar difficulties for the +R group and are clearly quite easy for these students. These five objectives have difficulties of 90.2% or above. MI has a difficulty of 79.4 for the +R students and thus stands out from the other objectives.

For the -R students the objectives are not as homogeneous as they are for the +R students. All the objectives sharply differentiate between the two groups. The order or relative difficulty of the objectives for the +R and -R students is quite consistent. The easier objectives include DW, RE, and DE; the more difficult objectives include MI, IN, and AL. This clustering of objectives is particluarly clear for the -R students.

Grade 3

The grade 3 results (see Figure 6) show more homogeneity of objective difficulty for the +R students than the -R students. The ordering or relative difficulty of the objectives is consistent for the +R and -R students. For both groups DE, DW and RE constitute relatively easy objectives while

AL, IN and MI constitute relatively more difficult objectives. All the objectives sharply discriminate between the +R and -R students.

Grade 6

The analysis of grade 6 (See Figure 7) shows a striking clustering of objectives. RE, DE and DW constitute the clearly easier objectives and IN, AL and MI are clearly the more difficult objectives. This clustering is apparent for both the +R and -R students. The objectives sharply differentiate the +R and -R students. The performance of the +R students on the difficult objectives is higher than the performance of the -R students on the easier objectives.

At grades 6 and 8, there are four options for each test item and therefore the objective difficulty expected by chance is .25. The performance of the -R students on AL (31.6%), IN (35.6%) and MI (37.7%) is not at this level but it is low enough to warrant serious concern.

Grade 8

The analysis of grade 8 (see Figure 8) shows more homogeneity of objective difficulty for the +R students compared to the -R students. Also apparent is the clustering of RE, DE and DW as easy objectives and MI, IN and AL as more difficult objectives. The objectives sharply differentiate between the +R and -R students.

Trends Across Grades

At grades 1 through 3, as well as 6 and 8, there is a very high consistency in the difficulty order of the BSAP Reading objectives for the +R and -R students. That is, the objectives that are relatively easier for one group of students (like Decoding and Word Meaning) tend also to be the easier BSAP objectives for other students as well. Similarly, the objectives that are relatively more difficult for one group of students (like Inference) tend also to be the harder BSAP objectives for the other student group.

All of the BSAP Reading objectives sharply differentiate the +R and -R students. In fact, the lowest performance on an objective for the +R students at any given grade is superior to the highest performance of the -R students at that grade.

The Decoding and Word Meaning (DW), Reference Skills (RE), and Details (DE) objectives in Reading appear to form a cluster of relatively easier objectives. This pattern becomes more and more pronounced after grade 1, and is most obvious at grade 6. The DW, RE, and DE cluster of objectives represents a set of reading tasks that load heavily on students' vocabulary knowledge and their ability to locate and recognize explicit and literally stated information.

The Analysis of Literature (AL), Inference (IN), and Main Idea (MI) objectives in reading appear to form a cluster of particularly difficult objectives. This pattern becomes more clearly pronounced after grade 1, and is also most obvious at grade 6. The AL, IN, and MI cluster of objectives represents a wide range of higher-order reading and thinking skills

that load heavily on inferential comprehension, requiring students to extract implied information and central meaning that is not literally or explicitly stated. The average percentage of correct responses to the AL, IN, and MI cluster items for the -R students is below 50% at each of the BSAP grades, 1 through 8.

STUDENTS' PERFORMANCE ON THE READING SUBSKILLS AND ITEMS

The analyses of students' performance on the BSAP Reading objectives can be refined by examining students' performance on the subskills that make up each objective. The reading subskills within the BSAP objectives are as follows:

DW: Decoding and Word Meaning

- DW 1: Sight Recognition (tested at all grades)
- DW 2: Phonetic Decoding (tested at grades 1,2,3)
- DW 3: Contextual Word Meaning (tested at grades 2,3,6,8)
- DW 4: Structural Word Meaning (tested at grades 6,8)

DE: Details

No subskills (tested at all grades)

MI: Main Idea

- MI 1: Restated Main Idea (tested at grades 1,2,3)
- MI 2: Paraphrased Main Idea (tested at all grades)
- MI 3: Inferred Main Idea (tested at all grades)

RE: Reference Usage

- RE 1: Selecting a Reference Source (tested at all grades)
- RE 2: Using a Reference Source (tested at all grades)

IN: Inference

- IN 1: Making Comparisons (tested at all grades)
- IN 2: Determining Cause and Effect (tested at all grades)
- IN 3: Drawing Conclusions (tested at all grades)
- IN 4: Predicting Outcomes (tested at all grades)

AL: Analysis of Literature

- AL 1: Nature of Information (tested at all grades)
- AL 2: Structural Elements (tested at all grades)
- AL 3: Rhetorical Devices (tested at all grades)
- AL 4: Critical Analysis (tested at grade 8)

Several procedures were employed to examine students' performance on the BSAP subskills. Subskill difficulties were calculated by obtaining the average item difficulty for the items assessing each subskill from 1981 to 1985. Subskill difficulties were calculated for the +R students and the -R students. The discrimination power of each subskill was calculated by subtracting the subskill difficulty of the -R students from the subskill difficulty of the +R students.

An additional analysis involved a careful examination of the relatively easy and relatively difficult items within each subskill in an effort to identify specific characteristics of items that accounted for differences in difficulty within subskills. The number and complexity of words and sentences in the item, the use of pronouns, the use of familiar and unfamiliar words, and a variety of other syntactic, semantic and structural features of the items were examined.

The results of the subskill and item analyses are presented for grades 1, 2 and 3 and then for grades 6 and 8. The subskills and items are examined within each objective in the order DW, DE, MI, RE, IN and AL. The discussion of the subskills and items within each objective begins with a subskill overview. This Subskill Overview describes what is involved in each subskill and how the subskills are assessed. The subskill and item analyses are then presented, along with numerous example items and passages designed to clarify description and explanation of students' performance on the subskills. The reader should note that these examples are <u>not actual BSAP items</u>, and therefore <u>do not necessarily meet BSAP item specifications</u>. These example items are included for the sole purpose of illustrating the results of the analyses reported herein.

DECODING AND WORD MEANING

GRADES 1, 2, 3

Subskill Overview

Decoding and Word Meaning (DW) items assess students' vocabulary knowledge, their ability to use word recognition skills, and their ability to determine the meaning of unfamiliar words in context. This objective has three subskills; two are tested at grade 1 and all subskills are tested at grades 2 and 3.

1. Sight Recognition (DW 1). This subskill involves the recognition of word meaning and is assessed at all grade levels. At grade 1, DW 1 is assessed by presenting students a single target word in a box located above three pictures. The students' task is to match the teacher-pronounced word with the appropriate picture. Words representing each of the picture options usually begin with the same letter as the target word in the box. At grades 2 and 3, the assessment procedure is similar, except that the target words are no longer read to students. Instead target words are underlined and set in a "neutral" stimulus context that does not provide students clues to word meaning. All target words are from the grade level of the S.C. Word List being tested, while words in the "neutral context" are below the grade level being tested. The "neutral context" provides no information about the meaning of the word.

2. Phonetic Decoding (DW 2). In this subskill, the teacher reads a nonsense word to the students who must choose the printed word which matches the word the teacher has pronounced. At grades 1 and 2, the target nonsense word and incorrect options contain only one syllable, while at grade 3 they may contain one or two syllables. The number, position, and general complexity of the phonetic elements assessed increases across grades 1 through 3 (see Teaching and Testing our Basic Skills for a more complete discussion of eligible phonetic elements and patterns). These items directly assess students' auditory discrimination and their ability to match auditory elements to

visually presented letter combinations. DW 2 items also assume that students understand the concept of nonsense words, that students possess adequate auditory acuity, and that teachers correctly and properly pronounce the nonsense words.

3. Contextual Word Meaning (DW 3). This subskill is not assessed in grade 1 and assessed at grades 2 and 3 by presenting an underlined and unfamiliar target word in a one or two sentence stimulus context. The target word is presumed to be highly unfamiliar to students. The revealing context uses words below the grade level being assessed. Incorrect options are typically opposite meanings, words whose spellings are similar to the correct option, and/or words representing plausible meanings.

The context for the unfamiliar target word can be of three types: (1) synonym, (2) antonym, or (3) situational context. Use of synonym context relies on the recognition of similarities while use of antonym context relies on the recognition of contrasts. Synonym and antonym context structures generally use only one or two words to suggest the meanings of the unfamiliar target word. Situational context structures generally suggest target word meanings with complete phrases, clauses, and/or sentences. The phrases, clauses, and sentences found in situational context items modify, rename, or illustrate a defining characteristic of the unfamiliar target word.

Examples of each type of contextual word meaning item appear below:

Example 1: Synonym Context

Item Stem: "A strange smell wafted, or floated, over the school."

In this synonym example, the unfamiliar target word has a vocabulary equivalent within the sentence that renames the target word. This vocabulary equivalency is explicitly cued by the signal word "or."

Example 2: Antonym Context

Item Stem: "Bill was happy, but Mary was sullen."

In this antonym example the context provides a vocabulary opposite of the unfamiliar target word, and explicitly cues the contrastive relationship of the two words with the signal word "but."

Example 3: Situational Context

Item Stem: "The decrepit car was falling apart."

In the situational example above, the phrase "was falling apart" describes a condition related to the unfamiliar target word "decrepit," and possesses no explicit cueing.

Grade 1: Decoding and Word Meaning

Analysis of Subskills 1 and 2

The average proportion of grade 1 students who correctly answer Decoding and Word Meaning items is displayed in Table 9 for students who reach and do not reach grade level standard, along with the subskill discriminations.

AVERAGE ITEM DIFFICULTY AND ITEM DISCRIMINATION FOR FIRST GRADE +R AND -R STUDENTS ON DECODING AND WORD MEANING SUBSKILLS

TABLE 9

Subskill	Average Difficulty for +R Students	Average Difficulty for -R Students	Average Item Discrimination
DW 1	.97	.69	.28
DW 2	.97	.75	.22

The DW subskills at grade 1 are extremely easy for +R students and relatively easy for -R students. Both subskills differentiate moderately well between the +R and -R groups.

Sight Recognition (DW 1): Grade 1

The analyses of these data and test items suggest that the amount of experience students have had with the words being assessed is the major factor affecting item difficulty. The more difficult items focus on target vocabulary words that are at the appropriate grade level but likely appear less frequently in grade level instructional materials.

Students' familiarity with vocabulary appears to control item difficulty. The difficulty of the DW 1 items at grade 1 is not related to the number of syllables in the target word, the location of the target word in the stimulus material, or the number of words in the stimulus sentence.

Phonetic Decoding (DW 2): Grade 1

There is little variation in the difficulty of DW 2 items for +R students in grade 1. There is more variation in the performance of -R students on these items. There are one or two specific phonetic elements which are consistently difficult for the -R students; however, inspection of these relatively difficult items reveals no general trends. The homogeneity of students' performance in the DW 2 subskill also suggests little with regard to patterns of student performance.

Grade 2: Decoding and Word Meaning

Analysis of Subskills 1, 2, and 3

The average proportion of grade 2 students who correctly answer Decoding and Word Meaning items is displayed in Table 10 for students who reach and do not reach grade level standard, along with the subskill discriminations.

AVERAGE ITEM DIFFICULTY AND ITEM DISCRIMINATION FOR SECOND GRADE +R and -R STUDENTS ON DECODING AND WORD MEANING SUBSKILLS

TABLE 10

Subskill	Average Difficulty for +R Students	Average Difficulty for -R Students	Average Item Discrimination
DW 1	.97	.66	.31
DW 2	.99	.94	.06
DW 3	.88	.48	.40

The DW subskills for +R students at grade 2 tend to be very easy. The +R students perform extremely well on DW 1 and DW 2 items and exhibit somewhat lower performance on DW 3 items. The -R students perform extremely well on DW 2 items, relatively well on DW 1 items, but demonstrate low performance on DW 3 items. Both DW 1 and DW 3 items differentiate between the +R and -R students very well. DW 2 items tend not to discriminate between the +R and -R students at grade 2 since they are extremely easy for both groups.

Sight Recognition (DW 1): Grade 2

Analyses of test items and student performance information suggest that the amount of experience students have had with the words being tested is the major factor affecting item difficulty. Although all vocabulary words tested are at the appropriate grade level, the more difficult items in DW 1 seem to focus on words that likely appear less frequently in grade level

instructional materials.

Students' familiarity with target vocabulary words appears to be the major determiner of item difficulty. The difficulty of DW 1 items is not related to number of syllables int he target word the location of the target word in the stimulus material, or any features associated with the stimulus context in which the target word is set.

Phonetic Decoding (DW 2): Grade 2

DW 2 items at grade 2 are very easy for both +R and -R students, and therefore do not differentiate between the two groups at all. No trends or patterns are noted with respect to DW 2 items at grade 2.

Contextual Word Meaning (DW 3): Grade 2

The relatively easy DW 3 items at grade 2 very often contain one or more of the following features. First, a single word or word pair in the stimulus material provides a clear and explicit contextual clue to the meaning of the unfamiliar target word. "Synonym" and "antonym" DW 3 items frequently are of this type (See Examples 1 and 2 in the Subskill Overview section). Second, stimulus context material associated with easier DW 3 items often contains explicit or implicit cause and effect relationships. Third, easier DW 3 items also tend to have relatively implausible incorrect options, particularly in combination with one of the first two features.

In contrast, the relatively more difficult DW 3 items tend to contain one or more of the following features. First, more difficult DW 3 items often contain an uncued phrase or a clause which must be comprehended completely in order for students to have the contextual information needed to determine the meaning of the unfamiliar target word. This feature is most typical among the "situational" Contextual Word Meaning items, with the exception of those that communicate a causal relationship (See Example 3 in the Subskill Overview section). Second, the more difficult DW 3 items are characterized by words or phrases containing relevant contextual information which are neither restricted in their meaning nor direct in their cueing. These broadly interpreted and indirect phrases or words tend to have multiple-meaning words, and are not specific to the unfamiliar target word (see Example 4 below). Third, difficult DW 3 items can employ contrastive cueing of the unfamiliar target word in the stimulus context material (See Example 5 below).

Example 4: Unrestricted Situational Context

Item Stem: "Too much government <u>largess</u> makes taxpayers angry."

Note that in Example 4 above, the phrase "makes taxpayers angry" is unrestricted in the sense that many things make taxpayers angry, "largess" on the part of the government being only one.

Example 5: Negated Situational Context

Item Stem: "Jim is submissive and won't stand up for his rights."

In Example 5, the clause "won't stand up for his rights" describes the absence of a characteristic, which in turn is a negated attribute of the unfamiliar target word "submissive." Neither of the example contexts above possesses explicit cueing. All of these factors tend to increase DW 3 item difficulty.

Grade 3: Decoding and Word Meaning

Analysis of Subskills 1, 2, and 3

The average proportion of grade 3 students who correctly answer Decoding and Word Meaning items is displayed in Table 11 for students who reach and do not reach grade level standard, along with the subskill discriminations.

AVERAGE ITEM DIFFICULTY AND ITEM DISCRIMINATION FOR THIRD GRADE +R and -R STUDENTS ON DECODING AND WORD MEANING SUBSKILLS

TABLE 11

Subskill	Average Difficulty for +R Students	Average Difficulty for -R Students	Average Item Discrimination
DW 1	.88	.49	.39
DW 2	.97	.85	.12
DW 3	.94	.62	.32

The DW subskills tend to be very easy for +R students at grade 3; their performance is extremely high in DW 2 and DW 3, and slightly lower in DW 1. Similarly, -R students perform quite well in DW 2, moderately well in DW 3, and perform low in DW 1. Both DW 1 and DW 3 items tend to differentiate very well between the +R and -R students, while the DW 2 items do not differentiate well between the two groups because they are easy for both groups.

Sight Recognition (DW 1): Grade 3

On Decoding and Word Meaning, the poorest performance of both +R and -R students at grade 3 is seen in the subskill DW 1. Analyses of student performance information and the items themselves suggest that the amount of experience children have had with the target vocabulary words in DW 1 is the main factor affecting item difficulty. While all target words are at the appropriate grade level, the more difficult items in DW 1 focus on

target words that likely appear less frequently in grade level instructional material.

Phonetic Decoding (DW 2): Grade 3

There is little variation in item difficulty for +R students in DW 2 items at grade 3. The -R students demonstrate more variation in performance on these items, but much of this is attributable to their consistent difficulty with one particular phonetic element.

Contextual Word Meaning (DW 3): Grade 3

The relatively easy DW 3 items at grade 3 very often can be characterized by one or more of the following features. First, easy DW 3 items typically contain a word or word pair in the stimulus context that provides a clear and explicit contextual clue to the meaning of the unfamiliar target word in the same context. "Synonym" and "antonym" DW 3 items frequently possess this characteristic. Second, the stimulus context sentence(s) of the easier DW 3 items frequently contains explicit or implicit cause and effect relationships. Third, the easier DW 3 items also tend to have less plausible incorrect options, this characteristic frequently occuring in combination with one of the first two features.

In contrast, the more difficult DW 3 items at grade 3 can typically be characterized by one or more of the following features. First, such items frequently contain an uncued phrase or clause which must be read and comprehended in its entirety before students can gain sufficient contextual information to determine the meaning of the unfamiliar target word. This feature is most typical among the "situational" Contextual Word Meaning

items, with the exception of those that communicate a causal relationship. Second, words or phrases containing relevant contextual information which are neither restricted in their meaning nor direct in their cueing, characterize the more difficult DW 3 items. These broadly interpreted and indirect phrases or words tend to have multiple meanings, and are not specific to the unfamiliar target words (See Example 4 in the grade 2 discussion). Finally, difficult DW 3 items can employ in the stimulus context material contrastive cueing of the unfamiliar target word by the direct negation of a familiar word (see Example 5 in the grade 2 discussion).

Summary

Decoding and Word Meaning: Grades 1, 2, 3

DW 1: Sight Recognition

DW 2: Phonetic Decoding

DW 3: Contextual Word Meaning

The +R students across grades 1 through 3 perform very well in Sight Recognition (DW 1), while the -R students perform relatively well in DW 1 at grades 1 and 2 but lower at grade 3. The Sight Recognition items tend to differentiate the +R and -R students very well at all grades. The primary explanation for performance differences among +R and -R students in DW 1, as well as among different items, lies mainly with the amount of experience students have had with the vocabulary being tested. At all grades, the target vocabulary words are at the appropriate grade level, but more difficult items tend to focus on words that likely appear

relatively less frequently in grade level instructional material.

It should be noted that the number of words on the S. C. Word List for grades 1, 2 and 3 is 341, 439 and 708, respectively. The percentage increase from grade 1 to grade 2 is 28%, while the increase from grade 2 to grade 3 is 61%. It appears possible that there is a decreased opportunity to provide direct instruction in grade level vocabulary at successive grades. This may in part explain the DW 1 performance drop at grade 3. No other factors such as number of syllables in target words, target word location in stimulus material, and so on, were related to item difficulty. These observations also support the contention that students' familiarity with vocabulary being tested is the major determiner of item difficulty.

The +R students perform extremely well in Phonetic Decoding (DW 2) at grade 1 through 3; item difficulties for this group ranged from .97 to .99 at these grades. The -R students also perform relatively well in DW 2; their lowest performance was in grade 1 where the average item difficulty for the -R students was .75. By grade 2, however, -R students demonstrate performance levels in DW 2 similar to that of +R students and -R students perform very well at grade 3 with the exception of their substantial difficulty with one item. In summary, DW 2 tends to discriminate between the +R and -R students at grade 1 only moderately well, and not at all in grades 2 and 3.

It should be noted that at grades 1, 2 and 3, there are many phonetic elements that are eligible for testing, but only a small percentage of these phonetic elements is assessed. This fact, combined with students' high performance on DW 2 items, makes it very difficult to specify categories of phonetic elements in which students are particularly weak or strong.

The +R students perform moderately well at grade 2 and perform well at grade 3 on Contextual Word Meaning items (DW 3 is not assessed at grade 1). The -R students at grade 2 demonstrate low performance on DW 3 items and at grade 3 relatively good performance on DW 3 items. Contextual Word Meaning items at grades 2 and 3 tend to differentiate very well between +R and -R students.

The more difficult DW 3 items in grades 2 and 3 appear to possess an increased reading load that requires more inferential processing. This contrasts to the easier DW 3 items in which the relevant contextual information is typically contained in a single word or word pair, is explicit in nature, and/or directly cued by means of key word and signal word repetition between stimulus and item stem. Response patterns of -R students suggest further that they are not checking to see if their choice of correct answer makes sense when substituted into the context provided by the stimulus material.

DETAILS

GRADES 1, 2, 3

Objective Overview

Details (DE) items assess students' recognition and recall of specific and explicit information from a reading selection. Details is tested at all grade levels and is not divided into subskills.

Students' accuracy in comprehending details in a reading selection is assessed by use of a simple paragraph of stimulus material followed by a question about one of the details in the paragraph. The stimulus material varies in length between 3 to 8 sentences depending on the grade level. In grade 1, the stimulus is a simple descriptive paragraph.

Stimulus material in grade 2 may include a description of a set of directions. At grade 3, the stimulus material may include a description of a set of directions, or may also include a sequence of events. The sequence of events is made clear by signaling words like "first," "next," "then," "after," "later," "last," and "finally."

In grades 1 and 2, the questions in the stem deal with "who," "how much," "how many," "what," "which" and "where." In grade 3, "how and "why" questions are also used.

In grade 1, the teacher reads the question to the students. In grades 2 and 3, the teacher directs the students to "Read the story in the box. Answer each question after it."

Grade 1: Analysis of Details

The items measuring Details have an average difficulty of .79 for the +R students, making Details one of the more difficult BSAP objectives for these students. The students who do not reach grade level standard have great difficulty with these items. The average item difficulty for the -R students is .33. The discrimination power of these items is .45. There are no items with difficulties of .50 or above for the students who do not reach grade level standard. For these students, 15 of the 30 items used have difficulties of less than .33, the proportion of students expected to answer correctly by chance.

Details (DE): Grade 1

The performance of the low achieving students seems strongly affected by prior knowledge. These students appear to choose answers to the questions that might be correct in their general experience but which are not always correct according to the stimulus material.

The nature of the incorrect answer selected by the low achieving students, along with their overall low performance, suggests that at least some of these students may not understand the task. These students may not have reached the level at which they can read and understand the stimulus material and/or they may not understand that the questions, which they read or which the teacher reads to them, should be answered based upon the information contained in the stimulus material.

The type of item that some students might answer based on their general knowledge and not based on stimulus information is illustrated by the following example:

> Billy painted a picture of some fruit. He painted the bananas green, the apples yellow and the grapes purple.

1: What color did Billy paint the bananas?

(A) Green [Stimulus Information]
(B) Yellow [General Knowledge]
(C) Purple

2: What color did Billy paint the grapes?

(A) Green
(B) Yellow
(C) Purple [Stimulus Information and General Knowledge]

Students' confusion about differences between stimulus based information and their own general knowledge about passage topics may lead to incorrect answers for question #1 above because students "know" that bananas are yellow. Student may answer question #2 correctly, however, through general knowledge alone, independent of stimulus information.

For all students, item difficulty is affected by the location in the story of the sentence needed to answer the question. Questions in which the required information is at the beginning of the story are easiest, followed by end-of-story information, and then medial information.

Questions dealing with "when" are most difficult for all students regardless of location of information. "Who" questions in the first position are easiest. "Where," "who," "what," and "how many" questions do not differ greatly in their difficulty.

Grade 2: Analysis of Details

Items measuring Details at grade 2 have an average difficulty of .93 for the +R students, and an average difficulty of .52 for the -R students. The discrimination power of these items is .41.

Details (DE): Grade 2

The performance of low achieving students is negatively affected by the degree of explicitness and specificity with which the detail is stated. The use of pronouns, in particular, increases the difficulty and discrimination of these items. Items in which the pronoun and referent are clearly associated in a single sentence are easier than items in which the pronoun and correct referent must be combined from two sentences. Questions

requiring comprehension of a sentence with two pronouns are especially difficult. For example, if the critical sentence is of the form, "She told him to go home" (pronoun-verb-pronoun action), questions like, "Where did Mary tell Bill to go?" are extremely difficult for the students who do not reach grade level standard.

Another type of implicit referencing also increases item difficulty. The use of kinship relations (e.g., brother, sister, father, mother, grandmother, grandfather) seems to introduce a difficulty factor similar to the use of pronouns.

The number of characters or agents in the stimulus material appears to increase item difficulty. The use of 2, 3 or 4 characters or agents, combined with pronoun referencing and other implicit references to the characters, increases difficulty to near the chance level for the low achieving students.

Items which are answered from explicitly stated specific details are clearly easier than items requiring referent re-coding for two or more characters or agents. The first type of item seems to assess comprehension of isolated literal details, while the second type seems to assess slightly more complex reading and language comprehension processes.

Questions dealing with "where," "who," or "where-when" are relatively easy. "What," "which," and "how" questions are more difficult. Questions in which the relevant information is at the beginning of the stimulus material are slightly easier than questions in which the relevant information is at the end of the stimulus material. In turn, questions in which the relevant information is at the end of the stimulus material are

slightly easier than the questions in which the relevant information is in the middle of the stimulus material. The order, from easy to hard, is beginning, end and middle.

Grade 3: Analysis of Details

For the +R third graders the items assessing Details have an average difficulty of .94. For students who have not reached grade level standard, the average item difficulty is .61. The discrimination power is .33.

Details (DE): Grade 3

There appears to be three major clusterings of items for Details in Grade 3. The first group of items contains, "who," "what," "why," how," "when," and "how often" questions. Items in this group have average difficulties of .88. The second group of items focuses on a sequence of events in the stimulus material and asks questions about what event happened in a particular sequence position. The sequence positions are explicitly cued in the stimulus material and in the stem of the questions with words like "first," "before," "after," "next," or "last." Thus, these sequence of events items assess recall or recognition of literally stated explicit details, like the first group of items. The average difficulty of these items is .89.

Like the second group, the third group of items deals with a sequence of events in the stimulus material and asks questions about what events happened in certain sequence positions. But the sequence positions of these events are not cued by temporal adverbs, adjectives, or prepositions

and, thus, these questions are not assessing recall or recognition of literally stated explicit details. Instead, these items require a more complex level of comprehension processing in which students must follow the inter-sentence sequence of events without explicit cues. The average difficulty of this third type of item is .80.

Examples of these three types of third grade Detail items are illustrated below with the following paragraph and associated questions.

> Sally went to the park to play. She saw Billy playing baseball. Then she went to the swimming pool. Later, Sally met Mary and they went to the zoo.

1: *Who played baseball? (Isolated detail question)*

2: *What did Sally do later? (Cued "what/sequence" question)*

3: *What did Sally do just before she met Mary? (Uncued "what/sequence" question)*

The students who do not reach grade level standard are more sharply differentiated from the other students by the uncued sequence of events items than by the other types of items. The position in the stimulus of the relevant information for answering the question does not affect item difficulties.

Summary

Details: Grades 1, 2, 3

In grade 1, item analyses for DE suggest that students, particularly those not reaching grade level standard, often do not understand the task required by these items. This confusion seems to arise because students do not clearly understand the difference between questions that can be answered based on their general knowledge and questions that can only be answered using information from a stimulus. Items in Details always require students to get information from a stimulus, but in some instances students tend to rely on their general knowledge of the world to answer questions. For first grade students, questions concerning "when" an event occurred were also particularly difficult.

In general, at grade 2, Detail items associated with stimulus selections that introduce two or more characters and use pronoun referencing are more difficult than other types of items. Item difficulty increases greatly when pronoun referencing of story characters extends across two or more sentences. This effect is more pronounced for students not reaching grade level standard.

Item analyses in Grade 3 Details suggest that a particular type of sequence of event item poses difficulty for students. Sequence of events questions that do not explicitly repeat signal word cues (e.g., "first," "last," "after," etc.) from the stimulus selection are clearly more difficult than other types of Detail questions.

MAIN IDEA

GRADES 1, 2, 3

Subskill Overview

For Main Idea, the student is presented with a reading selection from which the student must extract the central meaning or main idea. The student must identify three types of main ideas:

1. Restated Main Idea (MI 1). The student must identify a main idea which is restated as an answer option exactly as it appears in the reading selection. The main idea is explicitly stated in the reading selection as one sentence of a simple paragraph. The correct answer is a verbatim restatement of the main idea. MI 1 is assessed at grades 1, 2 and 3 only.

2. Paraphrased Main Idea (MI 2). The student must identify the main idea which is paraphrased from the reading selection. The main idea is explicitly stated in the selection, then reworded in the answer options. The degree of paraphrase increases as the grade level increases.

3. Inferred Main Idea (MI 3). The student must identify the main idea which is inferred from the reading selection. The main idea is not explicitly stated in the reading selection. As the grade level increases, the degree of inference increases.

Grade 1: Main Idea Subskills

Analysis of Subskills 1, 2, and 3

The average proportion of grade 1 students who correctly answer Main Idea items is displayed in Table 12 for students who reach and do not reach grade level standard, along with the subskill discriminations.

AVERAGE ITEM DIFFICULTY AND ITEM DISCRIMINATION FOR FIRST GRADE +R and -R STUDENTS ON MAIN IDEA SUBSKILLS

TABLE 12

Subskill	Average Difficulty for +R Students	Average Difficulty for -R Students	Average Item Discrimination
MI 1	.90	.45	.45
MI 2	.92	.50	.42
MI 3	.91	.45	.46

The MI subskills are all easy for the +R students, rather difficult for the -R students, and very powerful in differentiating between the two groups.

Restated Main Idea (MI 1): Grade 1

In Restated Main Idea (MI 1), several features seem to affect item difficulties. Relatively easy items have the main idea as the first sentence of the reading selection and thus correct answers repeat the first sentence in the stimulus material. Relatively easy items also tend to have incorrect options that repeat no more than one of the key words from the stimulus. More difficult MI 1 items tend to have the main idea as the last sentence of the reading selection and thus correct answers repeat the last sentence in the stimulus passage. Relatively difficult items also tend to have incorrect options that repeat more than one key word from the stimulus. In addition, relatively difficult MI 1 items have popular incorrect options

which are unrelated to the stimulus but are clearly true in the general experience of first graders.

Paraphrased Main Idea (MI 2): Grade 1

Two factors seem to affect the difficulty of the Paraphrased Main Idea (MI 2) questions. First is the number of sentences supporting the main idea. Second is the nature of the support the sentence(s) give to the main idea. Sentences in the stimulus material vary from providing direct or explicit support to providing indirect or implicit support. Indirect or implicit support requires students to make some degree of inference.

For the relatively easy items all or most of the sentences support the main idea and the support is direct and explicit. For the relatively difficult items, only one or two of the sentences provide support of the main idea and the support is indirect or implicit. For such items, students have few sentences cueing the main idea and the cues are remote and may require some inferential processing.

An example of an MI 2 item in which each stimulus sentence supports the main idea (MI 2) is shown below.

Example 1

Sally can make dresses. She can make coats and gloves, too. Sally can make many clothes.

What is the story about?

(A) The coat is old.

(B) The dresses and shoes are pretty.

**(C) Sally knows how to make many clothes.*

An example of an MI 2 item in which stimulus sentences provide mainly indirect support of main idea is shown below.

Example 2

Mother and I baked a cake. It was my birthday. It was fun to eat the cake.

What is the story about?

(A) The cat sleeps in the box.

**(B) We baked a cake.*

(C) Going to a party is fun.

Note that in the second example above, response option "C" is a more generalized statement, related to the passage, which tends to be true in the experience of most first graders. Options of this type are popular, particularly among -R students.

Inferred Main Idea (MI 3): Grade 1

Students' performance on Inferred Main Idea (MI 3) items is very similar to their performance on Restated Main Idea (MI 1). Two characteristics seem to make the MI 3 items particularly difficult for the students not reaching grade level standard. First, these students appear to be relying heavily on simple word or phrase associations between the stimulus material and the options. This simple paired-associative recognition strategy may lead to a correct answer, but more commonly leads to the selection of unrelated options that happen to have a verbal association with the stimulus material. Second, the students not reaching grade level standard often select options which are a general statement and thus resemble a main idea. These choices are true in the children's general experience, although unrelated to the stimulus material and unresponsive to the task demands of the items.

Grade 2: Main Idea Subskills 1, 2, 3

Analysis of Subskills 1, 2, and 3

The average proportion of grade 2 students who correctly answer Main Idea items is displayed in Table 13 for students who reach and do not reach grade level standard, along with the subskill discriminations.

AVERAGE ITEM DIFFICULTY AND ITEM DISCRIMINATION FOR SECOND GRADE +R AND -R STUDENTS ON MAIN IDEA SUBSKILLS

TABLE 13

Subskill	Average Difficulty for +R Students	Average Difficulty for -R Students	Average Item Discrimination
MI 1	.86	.61	.25
MI 2	.80	.40	.40
MI 3	.73	.33	.40

For both the +R and -R students, the Restated Main Idea is easiest, Paraphrased Main Idea is in the middle of the difficulty ordering, and Inferred Main Idea is the most difficult subskill. Restated MI items discriminate less sharply (.25) than Paraphrased and Inferred which have equivalent and very strong discrimination power (.40).

Restated Main Idea (MI 1): Grade 2

In Restated Main Idea, the easier items are those with stimulus material that has a topic sentence which is the obvious main idea. The stimulus material for such items generally has a single topic sentence directly supported by all the other sentences. For example:

Example 3

> It has been a dry summer. The grass is brown. The flowers have died. The small ponds have no water in them.

Students find it relatively simple to find the main ideas for such paragraphs. Identifying this type of the Restated MI is a simple recognition task.

More difficult MI 1 items have stimulus material in which the main idea is not explicitly elaborated upon or supported by each sentence. Students likely have more difficulty identifying the main idea of such material and, therefore, have more difficulty identifying the restatement of the main idea. This is particularly clear when the main idea is also a conclusion drawn from the other sentences. When the main idea is a conclusion, it is inferentially related to the other sentences as well as restated.

The analyses of these items suggest students may be answering the questions using two approaches, either separately or in combination. One approach involves inferring the main idea while reading the stimulus material and then looking for the main idea among the options. The second

approach involves the visual matching of sentences in the stimulus material with sentences in the options. The lower ability students may be using this second approach and thus improving their performance on MI 1, relative to MI 2 and MI 3. Only phonetic decoding (DW 2) and sight recognition (DW 1) are easier than MI 1 for the -R students at grade 2. This is consistent with the suggestion that MI 1 is more a sight recognition than a comprehension task for the -R students at grade 2.

The +R students may be using both approaches. While MI 1 is not an inferential task, students may infer the main idea while reading the stimulus material and then look for an option consistent with their inference. If they find one, they have, in effect, cross checked their answers. If they do not find an option which resembles what they infer to be the main idea, they may reread the stimulus and use the sight recognition strategy.

Paraphrased Main Idea (MI 2): Grade 2

The clarity with which a stimulus selection communicates the main idea affects the difficulty level of Paraphrased Main Idea items at grade 2. Relatively easy MI 2 items tend to have reading selections that possess one clear topic sentence, accompanied by sentences that provide direct support for the main idea. More difficult MI 2 items are typically accompanied by reading selections that possess two sentences that could be interpreted as topic sentences. Often, one sentence describes a general character preference or attitude. For example:

Example 4

> Mike studied a long time last weekend. He reviewed math on Friday. Mike studied French on Saturday. On Sunday, he studied science. Mike wanted to be a good student.

What is the story about?

(A) Mike was tired from studying.

(B) Mike studied French on Saturday.

**(C) Mike tried to get good grades.*

In such a selection, students appear to have difficulty determining whether the main idea relates to Mike's desire to be a good student or Mike's hard work at his studies. These are actually causally related in that "Mike works hard" because "Mike wants to be a good student." Options that restate a detail from the selection (e.g., on Saturday Mike studied French) are very popular incorrect answers for this type of item. A plausible explanation for the difficulty of these items suggests that students think the main idea deals with "Mike studied hard." Since they do not find this as one of the answer options, they choose as their answer a verifiable detail statement they know is accurate. This strategy is consistent with the earlier suggestion the MI 1 is a sight recognition task for the -R students. These students seem to be persisting in the use of the same strategy for identifying Paraphrased Main Ideas. The "sight recognition" strategy works for MI 1 items but not for MI 2 items, thus accounting for part of the increased difficulty of the MI 2 subskill.

A general cognitive problem is suggested by the performance of the -R students on the MI 2 items. The -R students do not seem to be able to deal with superordinate topics if the relationship is not explicitly cued by the stimulus material. For example, consider the revised stimulus selection in the box below.

Example 5

Mike wanted to be a good student. Mike studied a long time last weekend. He reviewed math on Friday. Mike studied French on Saturday. On Sunday he studied science.

The item is <u>easier</u> for the -R students with this revision because the relationship between the two general statements is cued by the ordering and contiguity of the sentences. The -R student can more easily see that the first two sentences are equivalent to:

Example 6

Mike wanted to be a good student so he studied a long time last weekend.

The original example obscures the superordinate-subordinate relationship of the two sentences (1) by using one as the first sentence and one as the last sentence and (2) by omitting signal words like "so,"

"therefore," or "because." The +R students can infer the relationship between the two general statements much more readily than the -R students.

Inferred Main Idea (MI 3): Grade 2

Relatively easy MI 3 items at grade 2 are cued by one sentence from the reading selection. As such, they do not necessarily require inter-sentence comprehension and inference. For such items, the degree of inference is minimal since only word or phrase substitution within one sentence is required.

More difficult MI 3 items in grade 2 often require students to make causal inferences among several events in the selection in order to correctly identify the inferred main idea. As such, these more difficult items tend to demand increased inferential processing by students. These items require students first to draw one or two inferences from the selections and then students must combine elements from these inferences into an accurate inferred main idea statement. For example, the following selection and accompanying item demonstrate this multiple inference process:

Example 7

> Last night Joe visited his friend Mike. The house was dark when Joe got there. When Joe went inside the house Mike turned on the light. Then, all of Joe's friends shouted "Happy Birthday."

Which sentence tells the story best?

*(A) *Mike and his friends had a surprise party for Joe.*

(B) Last night Joe visited his friend Mike.

(C) There are many parties at Mike's house.

In the above example, the correct inferred main idea is "A." To get the main idea, students must first infer that a surprise is involved based on the darkened house and friends shouting when the lights are turned on. They then must infer that, friends shouting "Happy birthday," implies a party. Finally, students must connect these two inferences to arrive at the inferred main idea involving a surprise party for Joe. Inter-sentence comprehension is needed to correctly identify the inferred main idea of this selection.

Grade 3: Main Idea

Analysis of Subskills 1, 2, and 3

The average proportion of grade 3 students who correctly answer Main Idea items is displayed in Table 14 for students who reach and do not

reach grade level standard, along with the subskill discriminations.

AVERAGE ITEM DIFFICULTY AND ITEM DISCRIMINATION FOR THIRD GRADE +R AND -R STUDENTS ON MAIN IDEA SUBSKILLS

TABLE 14

Subskill	Average Difficulty for +R Students	Average Difficulty for -R Students	Average Item Discrimination
MI 1	.85	.55	.30
MI 2	.81	.38	.43
MI 3	.80	.38	.42

As was the case in grades 1 and 2, the difference between the +R and -R students is less for MI 1 than for MI 2 or MI 3. The differences among MI 1, MI 2 and MI 3 are fairly small for the +R students. For the -R students, MI 2 and MI 3 have identical difficulties and are quite difficult (.38). On a 3-option multiple choice test this is very close to the chance level of .33. MI 2 and MI 3 are highly discriminating. Twice as large a proportion of +R students answer these items correctly in contrast to the -R students.

Restated Main Idea (MI 1): Grade 3

The relative difficulty of Restated Main Idea (MI 1) items is influenced by several factors. Items for which the main idea is the first or last sentence in the stimulus material are easier than items for which the main

idea comes from elsewhere in the stimulus. Easier items tend to have one wrong answer that very few students select. In contrast to the other two options, this unpopular option does not repeat a key word or key phrase from the stimulus material. Lastly, relatively easy MI 1 items tend to have simple, clear main ideas that are explicitly supported by several sentences in the stimulus material. More difficult MI 1 items have main ideas that are not as clear or well supported by the rest of the stimulus material.

Paraphrased Main Idea (MI 2): Grade 3

The Paraphrased Main Idea (MI 2) items do not vary greatly in their difficulties. The major factor affecting difficulty is the clarity of the main idea in the stimulus material. Items which have a clear, well cued and well supported main idea tend to be easy. Items for which the main idea requires combining information from several sentences without clear cues or support by simple elaboration are more difficult. The nature and degree of paraphrasing does not seem related to item difficulty.

On the more difficult MI 2 items, students have a strong tendency to choose a verbatim detail or a rephrased detail from the selection. These types of options often attract a larger proportion of -R students than the correct answer.

Inferred Main Idea (MI 3): Grade 3

Relatively easy MI 3 items have a simple main idea that is supported clearly and explicitly by several sentences in the stimulus. These relatively easy items often do not have an explicitly stated general topic

sentence.

The more difficult MI 3 items have a characteristic structure. These items have a correct answer which is a valid, generalized inference based on the reading selection. This "correct answer" is the "best" main idea from among the options listed but this "best" main idea from among the options listed is not always the most obvious main idea for the stimulus. Often the most obvious main idea from the stimulus is not listed among the options. In many cases, the most obvious main idea has an uncued subordinate or superordinate relationship with the option keyed as the correct answer. Students who inferred the main idea from the stimulus as they read may have a problem because they do not find their initial choice for main idea among the options. In such cases, students have a strong tendency to select a verbatim or restated detail as the correct answer.

Summary

Main Idea: Grades 1, 2, 3

MI 1: Restated Main Idea

MI 2: Paraphrased Main Idea

MI 3: Inferred Main Idea

Across all Main Idea subskills at grade 1, vocabulary repetition between the stimulus selection and the items appears to affect difficulty. Repetition of key vocabulary terms in incorrect options tends to make those options very attractive to low achieving students and thus increases the difficulty of the item. Repetition of only one or of no key words between stimulus and options tends to make those options less attractive. Overall, students appear to be relying heavily on the recognition of word asso-

ciations rather than on the understanding of main idea.

Grade 1 students may be experiencing some confusion in the task of main idea, particularly in Paraphrased and Inferred Main Ideas. Often, -R students tend to choose incorrect options that are like "main ideas" in that they are more generalized statements. Also, -R students appear to select their answers based on their prior knowledge and experience (e.g., bananas are yellow), not based on the stimulus material.

Finally, the difficulty of Paraphrased and Inferred Main Ideas items at grade 1, is affected by the degree and quality of main idea support present in the stimulus selections. Items with stimulus material that directly and explicitly supports the main idea in every sentence of the selection tend to be relatively easy. Items with reading selections that provide quantitatively less support for the main idea and/or support of a more implicit nature, tend to be more difficult because they require more inference skill.

At grade 2, the relative difficulty of Restated Main Idea items depends on the clarity and simplicity with which the main idea is stated. Easier items contain a single general topic sentence elaborated upon by each of the other sentences. MI 1 items are easy for all +R students and, compared to other subskills, relatively easy for the -R students who can respond to MI 1 items as sight recognition items.

The difficulty of Paraphrased Main Idea items at grade 2 also depends on the clarity with which the main idea is stated. The more difficult MI 2 items have stimulus selections with more than one general sentence. This general sentence subordinates the remaining details and supporting sentences.

The difficulty of Inferred Main Idea items at grade 2 is affected pri-

marily by the number and nature of the inferences students must make to answer the question correctly. More difficult MI 3 items require students to draw inter-sentence inferences from selections, then combine these inferences to form accurate inferred main idea statements.

At grade 3, the items measuring Main Idea subskills tend to be moderately difficult for the +R group, very difficult for the -R group and highly discriminating. In Restated Main Idea (MI 1) at grade 3, the degree of repetition of key vocabulary within the selections and item options appears to positively influence students' performance. Additionally, students' performance goes up in this subskill when the main idea sentence appears as the first or last sentence in the reading selection and when other stimulus sentences provide clear and explicit support for the main idea.

The difficulty of Paraphrased Main Idea (MI 2) items at grade 3 is affected by the clarity, amount of cueing, and selection support for the main idea. The type and number of words paraphrased does not affect item difficulty as much as the amount of inter-sentence inferencing. The -R students who incorrectly answer MI 2 items tend to choose incorrect options that repeat literal information from the selection. The difficulty of Inferred Main Idea (MI 3) items is also affected by the clarity of the main idea as a function of text structure and degree of inter-sentence inferencing required. The -R students operate only slightly above chance level on MI 2 and MI 3 items.

REFERENCE USAGE

GRADES 1, 2, 3

Subskill Overview

For Reference Usage (RE), students are required to select or use a variety of reference sources. As grade level increases, the number of reference sources for which students are responsible increases. In addition, the level of information detail requested from each reference source becomes increasingly more complex as grade level increases. Types of reference sources and their specific nature are the primary determinants of difficulty. The two subskills in Reference Usage are:

1. Selecting a Reference Source (RE 1). The student is presented a hypothetical reference need (e.g., where would you find the meaning of a word?). Then the student must select the correct reference source form the reference sources presented. The reference sources from which students must choose are either pictured (grades 1 and 2) or named (grade 3).

2. Using a Reference Source (RE 2). The student is presented with a reference source or an excerpt from a reference source, then asked to extract specific and explicit information asked for in the test item.

Eligible reference sources for both RE 1 and RE 2 at grade 1 include tables of contents, picture dictionaries, and calendars, and at grade 2 expands to include maps and glossaries. At grade 3, all of the foregoing

reference sources are eligible for assessment as well as encyclopedias, card catalogs, charts, graphs, and tables.

Grade 1: Reference Usage

Analysis of Subskills 1 and 2

The average proportion of grade 1 students who correctly answer Reference Usage items is displayed in Table 15 for students who reach and do not reach grade level standard, along with the subskill discriminations.

AVERAGE ITEM DIFFICULTY AND ITEM DISCRIMINATION FOR FIRST GRADE +R AND -R STUDENTS ON REFERENCE USAGE SUBSKILLS

TABLE 15

Subskill	Average Difficulty for +R Students	Average Difficulty for -R Students	Average Item Discrimination
RE 1	.88	.60	.28
RE 2	.93	.53	.40

The RE subskills for +R groups are easy, and for -R groups moderately difficult at grade 1. Both subskills tend to differentiate well between the +R and -R groups, with RE 2 being the more powerfully discriminating subskill.

Selecting a Reference Source (RE 1): Grade 1

The difficulty of RE 1 items at grade 1 is directly related to the type of reference source tested. In these items, students are given a reference need or requirement, then presented pictured versions of the possible sources that might fulfill the reference need. Easier items contain reference needs in the item stems that focus on calendars, while harder items focus on tables of contents and picture dictionaries.

There are two plausible explanations for students' lower performance on the table of contents and picture dictionary items. First, students' lack of familiarity with these reference sources and the kinds of information they contain might explain the difficulty of these items. Second, students may be visually confusing tables of contents and picture dictionaries since they are visually similar. In any case, the confusion between tables of contents and picture dictionaries, for whatever reason, is very pronounced with the lower achieving students.

Using a Reference Source (RE 2): Grade 1

The difficulty of RE 2 items at grade 1 appears to be primarily related to the type of the information source tested. Items that require the use of picture dictionaries tend to be easier while items that require the use of tables of contents tend to be more difficult. This difference appears to be related to the nature of the skill being required of students. In picture dictionary items, students must engage in simple picture-to-word matching tasks. In table of contents items, students must locate specific page information, given story titles.

2.5
2.8
3.2
3.6
4.0
2.2
2.0
100 mm
1.0
1.1
1.25
1.4
1.6
1.8
ABCDEFGHIJKLMNOPQRSTUVWXYZ
abcdefghijklmnopqrstuvwxyz
1234567890
1.0 mm
1.5 mm
2.0 mm
A5

Grade 2: Reference Usage

Analysis of Subskills 1 and 2

The average proportion of grade 2 students who correctly answer Reference Usage items is displayed in Table 16 for students who reach and do not reach grade level standard, along with the subskill discriminations.

AVERAGE ITEM DIFFICULTY AND ITEM DISCRIMINATION FOR SECOND GRADE +R AND -R STUDENTS ON REFERENCE USAGE SUBSKILLS

TABLE 16

Subskill	Average Difficulty for +R Students	Average Difficulty for -R Students	Average Item Discrimination
RE 1	.92	.57	.35
RE 2	.91	.55	.36

The RE subskills are easy for the +R students and moderately difficult for the -R students at grade 2. Items in both subskills tend to differentiate well between the +R and -R groups.

Selecting a Reference Source (RE 1): Grade 2

The difficulty of RE 1 items at grade 2 appears to be mainly related to the type of information source required by the item. Easier items require students to identify maps as the appropriate reference source for a hypothetical information need. More difficult items require students to

identify a table of contents or a picture dictionary for a hypothetical information need.

Differences between the easy items that focus on maps and the harder items that focus on other reference sources suggest that students, particularly -R groups, are not familiar with the reference sources that contain word meanings or those that yield story locations. Students may confuse tables of contents and picture dictionaries because they are visually similar as pictured in RE 1 items, in addition to the fact that students are confused about the type of information each possesses. Additionally, the easier map items in RE 1 at grade 2 tend to present a specific and concrete information need (e.g., "Where would you find the school?"), whereas the more difficult items involving picture dictionaries and tables of contents tend to present a more general and abstract information need (e.g., "Where would you look to find the meaning of a word?").

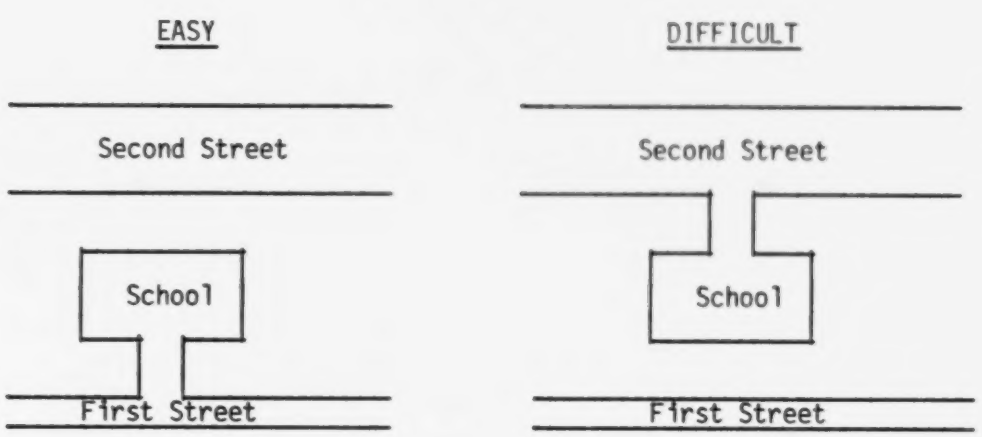
Using a Reference Source (RE 2): Grade 2

The type of reference source required by an item affects the difficulty of RE 2 items at grade 2. In general, table of contents items tend to be easier while picture dictionary items tend to be harder. The reference task required by these two types of items is somewhat similar. However, the table of contents items possess less total verbal information. Table of contents questions require students to first locate in the stimulus material a story title solicited in the stem, and then students must identify the page on which that story appears from among the options. The stimulus for picture dictionary items presents unfamiliar words to students in picture form with a verbal definition. Students must identify the

proper definition of the word from among the options.

Table of contents and picture dictionary items are answered with explicit information from the stimulus. But picture dictionary items require the stimulus-option matching of entire sentences, while the table of contents items require matching only one or two digit numbers. In both, the task is largely an explicit matching task, but more stimulus information must be matched in the picture dictionary items.

Certain types of map items tend to be easier than others. Items that require the extraction of information from maps tend to be easier when, for example, a building is located directly above the keyed street name. Structurally similar items become far more difficult for students when, for instance, a building is located directly below the keyed street name. In the following example, students might be asked "Which street is the school on?"



In each example above, students (particularly -R students) tend to answer "First Street." Two explanations for this response can be suggested. First, -R students may simply be unfamiliar with the meaning of the line

markers that connect the building to the street in a symbolic map representation. Second, -R students may be confused with the concept of "on" in this particular context; in a spatial sense, the school in both of the examples above is "on" (or directly above) First Street. This apparent confusion with "on" may be operating in conjunction with students' lack of familiarity with the line markers that connect buildings to streets in a symbolic map representation.

Grade 3: Reference Usage

Analysis of Subskills 1 and 2

The average proportion of grade 3 students who correctly answer Reference Usage items is displayed in Table 17 for students who reach and do not reach grade level standard, along with the subskill discriminations.

AVERAGE ITEM DIFFICULTY AND ITEM DISCRIMINATION FOR THIRD GRADE +R AND -R STUDENTS ON REFERENCE USAGE SUBSKILLS

TABLE 17

Subskill	Average Difficulty for +R Students	Average Difficulty for -R Students	Average Item Discrimination
RE 1	.88	.49	.39
RE 2	.91	.61	.30

The RE subskills for +R students are relatively easy, and for -R students moderately difficult at grade 3. Items in both subskills tend to differentiate well between +R and -R students.

Selecting a Reference Source (RE 1): Grade 3

The difficulty of RE 1 items at grade 3 is primarily related to the nature and type of reference source tested. In these items, students are given a specific reference need in the stem (e.g., "Where would you look to find out how to spell a word," "Where would you look to find out about the things a country produces," etc.), followed by options that name various reference sources.

Easier RE 1 items contain reference needs that require the use of dictionaries and maps. More difficult items in RE 1 focus on tables of contents and encyclopedias. In at least one instance, the inclusion of a proper place name in the stem of an item keyed to the encyclopedia as the correct answer seems to cause a large number of -R students to choose the incorrect option "map." No additional item features or student response patterns appear to account for variations in difficulty of items.

Grade 3 students in general and -R students in particular demonstrate a lack of familiarity with encyclopedias. On the other hand, all grade 3 students exhibit a high degree of familiarity with dictionaries.

Using a Reference Source (RE 2): Grade 3

The difficulty of RE 2 items at grade 3 is only slightly related to the type and nature of the reference source being tested. Easier items include charts, card catalogs, tables of contents, and maps and show little pattern with respect to type of reference source required. Harder items seem most often to focus on dictionaries and, to some extent, charts. However, the specific tasks being required of students within these items has far more

to do with item difficulty than the type of reference source being tested.

For example, there are very difficult and relatively easy RE 2 items requiring the use of charts. Easier chart items require students to read only across one row of information in order to obtain the fact or detail required to answer the question. Harder chart items require the same row reading strategy and also require students to read column headings at the same time in order to arrive at the fact or detail required to answer the question. In the task requiring both row and column reading, students appear to lose their place in the process of locating information from the chart since they tend to select incorrect options that contain a fact or detail that is physically near the correct answer in the stimulus chart.

Dictionary use items tend to be relatively difficult. These items present students with a dictionary entry containing multiple meanings for a word. The item stem then presents the same target word in a specific context and students must identify which of the several meanings given in the stimulus dictionary entry is appropriate to the context presented in the stem. Other dictionary use items present dictionary pages with guide words. These items require students to determine the page on which a word presented in the item stem would appear. These items range from moderately to very difficult.

At grade 3, students appear to have very little difficulty using card catalogs and tables of contents to find author, title, and other specific information. Map reading items in the third grade require students to determine the orientation of one location relative to another (e.g., "Which street is west of the hospital?"). These items are relatively easy for the +R student. However, the -R students have difficulty with these items when the target location is not lined up directly under the reference location

used to specify directional orientation. For example, -R students have trouble identifying the "school" as south of the "firehouse" if the school is shown as southwest or southeast of the firehouse.

Summary

Reference Usage: Grades 1,2,3

RE1: Selecting a Reference Source

RE2: Using a Reference Source

The RE subskills at grades 1 through 3 are among the easiest in BSAP relative to other subskills. The +R group at each grade performs very well in RE 1 and RE 2, while the -R group at each grade performs at or above the .50 level, for the most part. Items in RE tend to discriminate well at grades 1 through 3.

At grades 1 and 2, the difficulty of RE 1 and RE 2 items is primarily related to the particular reference sources used as stimulus material. Low achieving students at grades 1 and 2 appear to be unfamiliar with the nature of picture dictionaries and tables of contents, and they have difficulty getting specific information from these sources. Low achieving students at grade 2 perform relatively well in RE 2 items that use tables of contents, but have difficulty with RE 2 items that use picture dictionaries.

At grade 3, the influence of type of reference source is still strong in RE 1 items, but declines greatly in RE 2 items. Difficult RE 1 items at grade 3 focus on tables of contents and encyclopedias, while easy items focus on dictionaries and to some extent maps.

The difficulty of RE 2 items at grade 3 is primarily influenced by the complexity (rather than type) of the stimulus reference source and the specific nature of the skill and task demands. In general, items that require locating more information and information from several locations in the stimulus source tend to be more difficult. Items that require sentence comprehension (e.g., items that focus on dictionary usage) also tend to be more difficult.

INFERENCE

GRADES 1, 2, 3

Subskill Overview

For Inference (IN), students must read a stimulus selection and then determine the correct inference based on the selection. The student must make comparisons (IN 1), determine cause and effect (IN 2), draw conclusions (IN 3), or make predictions (IN 4). As grade level increases, the level of inference increases, the reading selections become longer, and the types of selections become more complex.

1. Making Comparisons (IN 1). The reading selections used in assessing this subskill in grades 1, 2 and 3 involve simple stories that describe some characteristic or quality of people or objects. Making comparisons (IN 1) assesses students' ability to compare persons or objectives relative to some characteristic or quality. In grade 1 a direct comparison between two persons or objects is required; in grades 2 and 3 a comparison of 3 persons or objects is required.

2. Determining Cause and Effect (IN 2). The reading selections used in assessing this subskill in grades 1, 2 and 3 involve simple stories that describe a series of events. At grades 2 and 3, the stimulus selections include description of the circumstances surrounding the events in the story. The item stem gives a "cause" and asks for an "effect" (outcome), or gives the "effect" and asks for the "cause." Both the cause and effect are explicitly stated in the stimulus material and the items assess students' ability to infer the relationship between the two.

3. Drawing Conclusions (IN 3). The reading selections used in assessing this subskill in grade 1 involve simple stories that describe an object, profession, time of day, or action. Grade 2 and grade 3 stimulus selections are simple stories that present a single topic developed with details. In grades 1 and 2, the conclusions students draw identify an object, location or action that can be inferred from the reading selection. In grade 3, the conclusions students draw identify general summary statements that can be inferred from the reading selection.

4. Predicting Outcomes (IN 4). In assessing this subskill, simple stories that describe a series of events are used at grade 1. At grades 2 and 3, the stimulus selections include description of the circumstances surrounding the events in the story. IN 4 items assess students' ability in predicting a reasonable action that could continue or conclude a series of actions or events. The stimulus selections present a description of a simple series of events (at grades 1 through 3) and/or the circumstances surrounding those events (at grades 2 and 3). IN 4 items at grades 1 through 3 require students to follow a series of events in order to predict a probable future action, attitude or feeling of one of the characters in the selection. At grade 3, IN 4 expands to include conditional information on the basis of which students must predict coming events, a character's future occupational preference, or a character's future feelings.

Grade 1: Inference

Analysis of Subskills 1, 2, 3, and 4

The average proportion of grade 1 students who correctly answer Inference items is displayed in Table 18 for students who reach and do not reach grade level standard, along with the subskill discriminations.

AVERAGE ITEM DIFFICULTY AND ITEM DISCRIMINATION FOR FIRST GRADE +R AND -R STUDENTS ON INFERENCE SUBSKILLS

TABLE 18

Subskill	Average Difficulty for +R Students	Average Difficulty for -R Students	Average Item Discrimination
IN 1	.83	.45	.38
IN 2	.87	.37	.50
IN 3	.85	.40	.45
IN 4	.87	.38	.49

The Inference subskills are very difficult for the -R students, moderately easy for the +R students, and sharply differentiate the two groups.

Making Comparisons (IN 1): Grade 1

The easier IN 1 items tend to explicitly cue the comparison with words like "very," "little," "all," "near," "far." In addition, easier items have reading selections that tend to have a characteristic structure. A first sentence contains a subject and two objects. The subject of the second sentence is second object of the first sentence and the subject of the third sentence is the first object of the first sentence. The comparison is made between the subjects of sentences 2 and 3 and requires reading only sentences 2 and 3. For example,

Example 1

I see a cat and a dog. The dog is far away by the yard. The cat is near me.

Which is nearer?

(A) dog

(B) yard

(C) cat

This type of item is relatively easy because (1) the contradictory cue "far away" is present, (2) the comparison involves only sentences 2 and 3, and (3) the signal word "near" is repeated between item stem and stimulus.

The more difficult IN 1 items do not contain explicit cue words. Furthermore, their syntactic structure is different. These selections have

three sentences, each with different subjects and objects. Thus, all three sentences must be considered in making the comparison. For example,

Example 2

Mary was talking. Joe was at a picnic. Jane shouts in the yard.

Who is more quiet?

(A) Mary

(B) Joe

(C) Jane

In example 2, students are not cued by an expression like "Mary was talking very quietly." Students must infer that a person talking is going to be "more quiet" than a person who is shouting.

Determining Cause and Effect (IN 2): Grade 1

The -R students have great difficulty with IN 2 items and the IN 2 items are highly discriminating. Two plausible explanations can be advanced for the difficulty and discrimination of the items assessing this subskill. First, a certain proportion of the -R students may not be able to read and comprehend the stimulus material even at the literal level. Second, even if some low achieving students can comprehend the stimulus material at the literal level, their response patterns suggest that they may not understand the nature and demands of items for which they must infer causes or effects from a reading selection.

Drawing Conclusions (IN 3): Grade 1

The major factor affecting IN 3 item difficulties at grade 1 is the use of pronouns and their antecedents in the stimulus material and item stems. Difficult items use extensive pronoun substitution as illustrated below:

Example 1

I have a pet. It can fly. It can sing.

What could it be?

Easier items tend to use more direct and concrete referencing of stimulus characters, as illustrated below:

Example 2

I have a pet. My pet can fly. My pet can sing.

What is my pet?

A second notable feature of students' responses to IN 3 items is the common choice of extremely implausible answers. In some instances, -R students choose answers that are contradicted three to four times by the stimulus information. The choice of contradicted options suggests that some first grade students may not understand the task.

Predicting Outcomes (IN 4): Grade 1

The response patterns of -R students to IN 4 items suggest that they may not be able to read and literally comprehend the stimulus selections. Further, they may not understand the nature and demands of the IN 4 items that require following an explicit series of events and then inferring a probable "next event" that is not explicitly stated.

Easier IN 4 items tend to have options that are implausible given the series of events. Often each sentence in the reading selections of the easier items cues the correct answer and thus students can answer the item correctly if they can comprehend a single sentence from the stimulus. In contrast, more difficult items require students to comprehend more events from among the sentences in the reading passage. Also, the more difficult items often contain wrong answers that are cued by means of key word association to the stimulus selections. Students operating at the literal level find these "cued" wrong answers appealing.

Grade 2: Inference

Analysis of Subskills 1, 2, 3, and 4

The average proportion of grade 2 students who correctly answer Inference items is displayed in Table 19 for students who reach and do not reach grade level standard, along with the subskill discriminations.

AVERAGE ITEM DIFFICULTY AND ITEM DISCRIMINATION FOR SECOND GRADE +R AND -R STUDENTS ON INFERENCE SUBSKILLS

TABLE 19

Subskill	Average Difficulty for +R Students	Average Difficulty for -R Students	Average Item Discrimination
IN 1	.94	.43	.51
IN 2	.93	.45	.48
IN 3	.92	.45	.47
IN 4	.86	.34	.52

The IN subskills are easy for the +R students, very difficult for the -R students, and highly discriminating.

Making Comparisons (IN 1): Grade 2

There are no specific content features of the IN 1 items that explain the variation in item difficulty. The homogeneity of item difficulty within the +R and -R groups and the high discrimination suggest that IN 1 items may be more sensitive to general reading ability than the specific features of the individual item.

Determining Cause and Effect (IN 2): Grade 2

At grade 2, -R students very likely have difficulty literally comprehending the stimulus material. At the same time, the response patterns of the -R students suggest that they may not fully understand the inference task involved. As with IN 1, there are no obvious content features of these items that account for variation in item difficulty.

Drawing Conclusions (IN 3): Grade 2

The most striking observations about students' performance in IN 3 at grade 2 is that the items are extremely easy for students reaching grade level standard and considerably more difficult for students who do not reach standard. At grade 2, the average item difficulty for students reaching standard is .93 and the average difficulty for those not reaching standard is .45.

Analyses of students' response patterns suggest that low achieving students tend to select wrong answers that repeat words from the stimulus selection. In addition, -R students appear to often use information contained in the beginning portions of the stimulus material in selecting

their answers. The correct answers usually require the use of the whole reading passage while many incorrect answers are related to the beginning portions of the material.

Predicting Outcomes (IN 4): Grade 2

At grade 2, -R students may have difficulty literally comprehending the stimulus selections and at the same time they may not fully understand the nature of the inference task required. Further, the incorrect options in IN 4 items are often incorrect because they are negated or contradicted in the stimulus selection. Easier items tend to have incorrect options that are literally and explicitly negated or contradicted in the stimulus selection. More difficult items have incorrect options that are implicitly negated or contradicted in the stimulus selection. Implicit negation requires considerable inferential processing by students and may explain why IN 4 items are so sharply discriminating.

Grade 3: Inference

Analysis of Subskills 1, 2, 3, and 4

The average proportion of grade 3 students who correctly answer Inference items is displayed in Table 20 for students who reach and do not reach grade level standard, along with the subskill discriminations.

AVERAGE ITEM DIFFICULTY AND ITEM DISCRIMINATION FOR THIRD GRADE +R AND -R STUDENTS ON INFERENCE SUBSKILLS

TABLE 20

Subskill	Average Difficulty for +R Students	Average Difficulty for -R Students	Average Item Discrimination
IN 1	.86	.51	.35
IN 2	.90	.43	.47
IN 3	.93	.49	.44
IN 4	.85	.38	.47

IN items at grade 3 are easy for +R, students, difficult for -R students, and differentiate the two groups very sharply.

Making Comparisons (IN 1): Grade 3

The relatively easy IN 1 items contain quantitative adjectives which cue the relevant dimension on which the comparsion is to be made (e.g., "farthest"). The relatively more difficult items do not tend to have the single word explicit cues. Rather, the more difficult items imply the relevant dimension in a phrase, clause or entire sentence.

Determing Cause and Effect (IN 2): Grade 3

At grade 3 there is some evidence that students are either unable to read the stimulus material or read quickly and incompletely. This is suggested by items in which a larger proportion of low achieving students

select answers that are explicitly contradicted in the stimulus than choose the correct answer.

Drawing Conclusions (IN 3): Grade 3

The most striking observation about students' performance in IN 3 at grade 3 is that the items are extremely easy for those students who reach grade level standard, and extremely difficult for students who do not. Analyses of grade 3 students' performance on IN 3 items suggest that -R students tend to choose incorrect options that repeat key vocabulary from the stimulus selection. In addition, these students often appear to be restricted to using information contained in the beginning portions of the stimulus selection in choosing wrong answers.

Predicting Outcomes (IN 4): Grade 3

At grade 3, the -R students may have difficulty literally comprehending the stimulus selections, and at the same time may not fully understand the nature of the inference task required in IN 4. Further, IN 4 items tend to possess incorrect options, the substance of which negates or contradicts information given in the stimulus. Easier items tend to have incorrect options of this sort that are literally or explicitly contradicted in the stimulus. More difficult items possess incorrect options that are implicitly negated or contradicted by the information in the selection, thus requiring more inferential processing by students.

Summary

Inference: Grades 1, 2, 3

IN 1: Making Comparisons

IN 2: Determining Cause and Effect

IN 3: Drawing Conclusions

IN 4: Predicting Outcomes

At grades 1 through 3, Inference items tend to range from moderately easy to very easy for +R students. Inference items range from moderately difficult to very difficult for -R students.

The most striking feature of the Inference subskills is how sharply and consistently they discriminate between the +R and -R students. The consistency of the discrimination across subskills and grades suggests that the +R and -R students are differentiated on the basis of some general inference ability or aptitude. It may be that the inference items separate -R students who read at only a literal, word recognition level from students (+R) who are capable of more abstract comprehension abilities. One would expect the difference between two such groups of students to be large and consistent no matter what type of inference item they answered.

Direct and explicit cueing in grade 1 Inference items has a large effect on item difficulty. In Making Comparisons (IN 1), selections and items that directly name or cue the dimension along which the comparison is to be made (e.g., "nearest," "farthest," etc.) tend to be easier. When students must infer both the dimension of comparison and the relative position of elements being compared, items are much more difficult.

The extent of inter-sentence comprehension an item requires also

affects its difficulty. In Predicting Outcomes (IN 4) items, easier items tend to describe a series of events, each of which cues the correct answer. More difficult IN 4 items require students to comprehend multiple events in the sequence in order to correctly identify a probable outcome.

Finally, at grade 1, items which include inter-sentence pronoun/antecedent substitutions tend to be more difficult. In Drawing Conclusions (IN 3), stimulus selections that name an agent or character and then subsequently refer to that agent in the selection and item stem with pronoun substitution tend to be more difficult than IN 3 items that do not use pronoun referencing.

Explicit cueing between stimulus selection and items via key word and/or signal word repetition affects the difficulty of Inference items at grade 2. Items are not made easier by the direct cueing of the correct answer. Rather the more difficult items, for example in Drawing Conclusions (IN 3), cue incorrect answers. Key vocabulary repetition in wrong options appears to draw students, particularly -R students, to those wrong answers.

The relative degree of inferencing required to correctly answer Inference items at grade 2 also influences item difficulty. In Predicting Outcomes (IN 4) items, the easier items typically possess incorrect options that are explicitly negated or contradicted in the selection. Harder IN 4 items tend to possess incorrect options that are implicitly contradicted in the accompanying selections, thus requiring students to make additional inferences in order to recognize these options as incorrect answers.

The difficulty of Inference items at grade 3 is strongly influenced by the type of cueing available to students. In Making Comparisons (IN 1), easier items typically name the dimension of comparison explicitly in one

word. More difficult IN 1 items tend to use a phrase or clause to communicate the dimension of comparison, and often the dimension is implied rather than explicitly named.

Vocabulary cueing of wrong answers is associated with more difficult items in Drawing Conclusions (IN 3). These items tend to possess incorrect options that repeat key words from the selection and draw more student responses particularly from among the -R students. Finally at grade 3, Predicting Outcomes (IN 4) items with incorrect options that are explicitly negated or contradicted in the reading selection tend to be easier than other items. Harder items often have incorrect options that are implicitly contradicted in the selection, thus requiring students to make inferences in order to eliminate these options as correct answers.

ANALYSIS OF LITERATURE

GRADES 1, 2, 3

Subskill Overview

In Analysis of Literature, students are presented with a reading selection which they must critically analyze. There are three AL subskills in grades 1, 2 and 3.

1. Nature of Information (AL 1). For this subskill students must identify something that is "make believe" from a story. The "make-believe" or fantasy element changes across the grades. In grades 1, 2 and 3 the fantasy element may describe an

action that goes beyond or violates a subject's natural capabilities, e.g., "the boy jumps over the house." In grades 2 and 3, the fantasy element may involve supernatural transformations, e.g., "The man's blue suit turned into white feathers." In grade 3, the fantasy element may include mythical or imaginary characters (e.g., giants).

2. Structural Elements (AL 2). For this subskill, students read a story and then must answer a question which asks them to identify the plot, a character, or the setting of the story. The AL 2 questions asking about the "plot" require students to identify a statement that accurately summarizes the story. The AL 2 items dealing with "character" generally ask students to identify a story character's personality features and/or abilities. These items typically require inference. The AL 2 items dealing with "setting" ask students to identify the place where the story occured (e.g., in the park) at grades 1 and 2. At grade 3, the AL 2 items may also require a detailed description of the environment (e.g., in an old house in the forest).

3. Rhetorical Devices (AL 3). In grades 1 and 2, items assessing this subskill require students to identify words that rhyme from a stimulus passage. The stimulus passage is a poem the students read. The rhyme is based on the repetition of vowel and consonant sounds at the ends of words (e.g., baboon, harpoon). In grade 3, AL 3 items require students to identify "which two things are being compared in the poem." The comparison pair is a simile marked by the word "like" or "as."

Grade 1: Analysis of Literature

Analysis of Subskills 1, 2, and 3

The average proportion of grade 1 students who correctly answer Analysis of Literature items is displayed in Table 21 for students who reach and do not reach grade level standard, along with the subskill discriminations.

AVERAGE ITEM DIFFICULTY AND ITEM DISCRIMINATION FOR FIRST GRADE +R AND -R STUDENTS ON ANALYSIS OF LITERATURE SUBSKILLS

TABLE 21

Subskill	Average Difficulty for +R Students	Average Difficulty for -R Students	Average Item Discrimination
AL 1	.76	.33	.43
AL 2	.81	.33	.48
AL 3	.73	.31	.42

Both the +R and -R students have difficulty with the Analysis of Literature subskills. For the +R students, AL 2 is slightly easier than AL 1 and AL 3, although the subskill difficulties are fairly homogeneous. The -R students appear to have no idea how to answer the AL questions. With three options per item, the item difficulties would be .33 if students were guessing at random. This is the level achieved by the -R students. The AL items are extremely discriminating. More than twice as large a proportion of +R students answer the AL items correctly as the -R students.

Nature of Information (AL 1): Grade 1

On AL 1 items, students must identify the option which is make-believe. Items involving people as story characters are slightly easier than items involving animals. The response patterns of the -R students and some of the +R students strongly suggest that they simply do not understand the task required by the item. They may not have the basic con-

cept of make-believe. In addition, they may not understand the words well enough to make the make-believe assessment even if they understood the concept. The -R students have a strong tendency to select incorrect options that are true or make sense in their own experience.

Structural Elements (AL 2): Grade 1

On Structural Elements (AL 2) items, students generally do better on items dealing with setting or plot than questions about characteristics of people in the story. The setting items can generally be answered by simple one-word associations. The correct answers to the plot items generally entail a minor literal paraphrase of a topic sentence/main idea. The character items are of the form, "What does the story tell about (name of person)?" Often, the story character in question is the subject of two or three sentences and students must make an inference or draw a conclusion based on this information. Thus, inter-sentence inferences are required in many AL 2 items, making them very difficult for the -R students. The -R students frequently pick incorrect options that repeat an easily recognized word or phrase from the stimulus. As seen in AL 1, many students may not understand what they are supposed to do in responding to AL 2 items.

Rhetorical Devices (AL 3): Grade 1

First grade students must identify from a poem a pair of words that rhyme in AL 3 items. The item structure generally includes an incorrect word pair that does not rhyme and an incorrect word pair that rhymes but which contains a word that is not in the poem. The AL 3 items are very

difficult for all students but are practically impossible for the -R students. The -R students may not understand the nature of the task at all. For several of the AL 3 items, about one third of the -R students select each of the options. Furthermore, -R students often select the first rhyming pair listed in the option set regardless of whether the word pair is taken from the poem, or they pick the more familiar and recognizable pair.

Grade 2: Analysis of Literature

Analysis of Subskills 1, 2, and 3

The average proportion of grade 2 students who correctly answer Analysis of Literature items is displayed in Table 22 for students who reach and do not reach grade level standard, along with the subskill discriminations.

AVERAGE ITEM DIFFICULTY AND ITEM DISCRIMINATION FOR SECOND GRADE +R AND -R STUDENTS ON ANALYSIS OF LITERATURE SUBSKILLS

TABLE 22

Subskill	Average Difficulty for +R Students	Average Difficulty for -R Students	Average Item Discrimination
AL 1	.93	.41	.52
AL 2	.94	.46	.48
AL 3	.84	.38	.46

For each of the subskills, twice as large a percentage of +R students answer correctly compared to the percentage of -R students. AL 1 and AL 2 items are quite easy for the +R students and AL 3 is the most difficult subskill for both the +R and -R second graders.

Nature of Information (AL 1): Grade 2

In the AL 1 items, students are required to identify the statement which is make-believe. The performance of the -R students suggests that some of them may not understand this task or may not posssess a clear concept of make-believe. The use of pronouns seems to increase item difficulty. Within the +R and -R groups, the difficulty level of the items is fairly homogeneous. Besides pronoun use, there are no obvious structural features of the items which influence their difficulty.

Structural Elements (AL 2): Grade 2

The AL 2 items require students to answer questions about plot, character, and setting. Items asking "What does the story tell us about Joey?" are character items and are more difficult than plot items (e.g., "What is this story about?") and setting items (e.g., "Where did the story take place?") The character items have pronouns in all the options and generally require students to make an inference or come to some conclusion about a person based on information from more than one sentence in the story. The plot and setting items are more literal, requiring little inferential processing and are easier than the character items.

Rhetorical Devices (AL 3): Grade 2

Rhetorical Devices (AL 3) is the most difficult subskill for both the +R and -R students. Some of the rhyming pairs the students select are simple and obvious rhymes but contain a word that is not in the poem. The more difficult rhymes are pairs of words that are not cued directly from their spelling. For example, "hair" and "where" are not cued by spelling, whereas "fall" and "ball" are clearly cued. The items cued by spelling pattern allow students to answer the questions by simple matching of letters; the uncued items cannot be answered in this way and clearly require students to know the pronunciation of more vocabulary words in order to correctly answer the items.

Grade 3: Analysis of Literature

Analysis of Subskills 1, 2, and 3

The average proportion of grade 3 students who correctly answer Analysis of Literature items is displayed in Table 23 for students who reach and do not reach grade level standard, along with the subskill discriminations.

AVERAGE ITEM DIFFICULTY AND ITEM DISCRIMINATION FOR THIRD GRADE +R AND -R STUDENTS ON ANALYSIS OF LITERATURE SUBSKILLS

TABLE 23

Subskill	Average Difficulty for +R Students	Average Difficulty for -R Students	Average Item Discrimination
AL 1	.98	.63	.35
AL 2	.96	.56	.40
AL 3	.51	.31	.20

AL 1 and AL 2 are very easy for the +R students and difficult for the -R students at grade 3. AL 3 in grade 3 is extremely difficult for both +R and -R students. The -R students operate below the level of chance performance. (By randomly picking answers, 33% of a group is likely to answer correctly but only 31.4% of the -R group answer AL 3 items correctly.)

Nature of Information (AL 1): Grade 3

The relatively difficult AL 1 items are cued as make-believe by the reference to a mythical character. If students are unfamiliar with the vocabulary term for the mythical being, they likely will not be able to identify the make-believe statement correctly. For example, "The elf sang to Bill" is a more difficult type of make-believe statement to identify than "The bed flew over the field." This easier type of item contains familiar objects/characters doing make-believe activities.

Structural Elements (AL 2): Grade 3

The AL 2 items assess students' understanding of plot, character and setting. These are very easy for the +R students. The more difficult items are plot items requiring inter-sentence comprehension, and possessing no simple vocabulary cueing. These more difficult AL 2 items are very similar in task demand to inferred main idea items and sharply differentiate +R from -R students.

Rhetorical Devices (AL3): Grade 3

Rhetorical Devices in (AL 3) grade 3 are designed to assess students' understanding of similes. Students must identify which two elements are compared in a poem. The students' extremely low performance suggests that they do not understand the demands of the task. The comparisons are cued by the words "as" and "like," and which of these two words is used does not affect item difficulty. The difficulty of AL 3 items is clearly revealed by the nature of the answers students choose. Students tend to choose incorrect options that are rhyming word pairs and/or pairs of words that are commonly associated. The selection of these kinds of incorrect options supports the suggestion that students do not understand the nature of the task. The few relatively easy AL 3 items contain options with no rhyming word pairs or no closely associated word pairs.

Summary

Analysis of Literature: Grades 1, 2, and 3

AL 1: Nature of Information

AL 2: Structural Elements

AL 3: Rhetorical Devices

The AL subskills are extremely difficult for -R students and moderately difficult for +R students at grade 1. At grade 2, they are extremely difficult for -R students, but relatively easy for +R students. In grade 3, AL 1 and AL 2 are very easy for +R and somewhat difficult for -R groups, while AL 3 is extremely difficult for all students. Items assessing all AL subskills at grades 1 and 2 are highly discriminating. At grade 3, items in AL 1 and AL 2 are also highly discriminating, while items in AL 3 are moderately discriminating.

In grade 1, -R students appear to be highly unfamiliar with the concept of make-believe, unable to make identifications about common story elements in stories, and unfamiliar with matching rhyming words. The +R students' performance is also relatively low in all AL subskills at grade 1. In AL 2, students tend to more often correctly identify make-believe elements when these elements involve people rather than animals. In AL 2, plot and setting items tend to be easier than items that deal with analysis of characters which require more inference and inter-sentence comprehension. AL 3 items require a rhyming task that appears unfamiliar to many students, but particularly to -R students.

In grade 2, all AL subskills are extremely difficult for -R students and quite easy for +R students. In AL 1, the -R students do not seem to

understand the concept of make believe, and may be unfamiliar with vocabulary words that cue such items (e.g., "giant"). In AL 2, the character analysis items tend to be more difficult than items requiring identification of plot or setting. This is probably because character items tend to require more inference and inter-sentence comprehension. AL 3 performance suggests that students are either unfamiliar with the concept of rhyming words, or they do not know the pronunciation of vocabulary words needed to identify the rhymed words.

In grade 3, AL 1 and AL 2 items are very easy for +R students and somewhat difficult for -R students. Both +R and -R students have great difficulty in identifying members of a comparison in the AL 3 subskill. AL 1 items that are cued as make believe by the presence of a mythical being tend to be more difficult. Difficult AL 2 items tend to require inter-sentence comprehension and possess little explicit cueing. The extremely low performance and option selections in AL 3 suggest that +R and -R students do not have a clear understanding of the skill and task demands of the items. (These items require students to identify two elements being compared in a poem, marked by "like" or "as".)

DECODING AND WORD MEANING

GRADES 6, 8

Subskill Overview

Decoding and Word Meaning is an objective designed to assess how well students can use word recognition skills and how well they can determine the meanings of words. Students' decoding and word meaning skills are assessed by three subskills in grades 6 and 8. The DW subskills are defined as follows:

1. Sight Recognition (DW 1). Sight Recognition items assess students' ability to recognize words by correctly identifying their meaning. Reading selections contain one or two sentences with minimal context clues. The target word is underlined in the sentence and comes from the grade level word list. The other words in the sentence are below grade level. The questions for DW 1 items read, "In this sentence, (target word) means:"

2. Contextual Word Meaning (DW 3). Contextual Word Meaning items are designed to assess how well students can determine the meaning of an unfamiliar word used in a selection from the context of the selection. The selection contains one or two sentences with one or more context clues for each selection. The target word is very unfamiliar and the other words in the sentence are below grade level. The questions for DW 3 items read, "In these sentences, the (target word) means:"

3. Structural Word Meaning (DW 4). Structural Word Meaning items are designed to assess how well students can determine the meaning of a word in a selection by using information about a prefix or suffix. The item selection contains one or two sentences with minimal context clues. The target word is unfamiliar to students and the other words in the selection are below grade level.

Grade 6: Decoding and Word Meaning

Analysis of Subskills 1, 3, and 4

The average proportion of grade 6 students who correctly answer Decoding and Word Meaning items is displayed in Table 24 for students who reach and do not reach grade level standard, along with subskill discriminations.

AVERAGE ITEM DIFFICULTY AND ITEM DISCRIMINATION FOR SIXTH GRADE +R and -R STUDENTS ON DECODING AND WORD MEANING SUBSKILLS

TABLE 24

Subskill	Average Difficulty for +R Students	Average Difficulty for -R Students	Average Item Discrimination
DW 1	.94	.65	.29
DW 3	.94	.72	.22
DW 4	.83	.52	.31

For the +R students, DW 1 and DW 3 are quite easy whereas DW 4 is clearly more difficult. DW 4 is also the most difficult subskill for the -R students. The -R students also find DW 1 to be somewhat difficult (.65). The DW 1 and DW 4 subskills differentiate the +R and -R students quite well at about .30.

Sight Recognition (DW 1): Grade 6

Students perform quite well on Sight Recognition items. Easier items appear to focus on vocabulary that likely occur frequently in students' instructional materials; more difficult items, while at the appropriate grade level, are likely to occur infrequently in students' reading materials. It is also likely that the more difficult items assess the recognition of words that occur infrequently in the oral language students hear and use. In some of the more difficult items, students select incorrect options which have a clear phonemic association with the target word.

Contextual Word Meaning (DW 3): Grade 6

The relatively easy DW 3 items are heavily cued in one of two ways. First, some easy items contain more than one signal word, each of which directly cues the correct answer and assists in eliminating the choice of wrong answers. Second, some easy DW 3 items have a meaningful prefix in the unfamilar target word which is repeated in the correct answer. Such items can be answered correctly by using a "match the salient prefix" strategy. More difficult DW 3 items provide cueing in multiple word phrases rather than in single words or prefixes. Item difficulty is compounded when the meaning of the target word is cued in a negative phrase with appropriate context and students must identify the meaning of the unnegated word. An example of this type of item is shown below:

The boat was not visible in the thick fog.
Even with binoculars, I could not see it.

In this sentence, visible means:

ans: able to be seen

The stimulus provides two cues in different phrases: 1) the boat is in the fog and 2) it could not be seen. The item asks for the meaning of "visible" and students must infer that since "not visible" means the boat "could not be seen" then "visible" must mean the opposite or is "able to be seen." This type of item is especially difficult for the -R students.

Structural Word Meaning (DW 4): Grade 6

The difficulty of DW 4 items is affected by several factors individually and in combination. First, target words with familiar root words are relatively easy. (This first factor is similar to the DW 1 dynamic in which more common and more frequently occurring words are found in the easier items.) Second, target words with a prefix which directly cues a word or phrase in the correct answer are also relatively easy. When this direct cueing is not present, items tend to be more difficult. Third, the more difficult DW 4 items contain a prefix which does not contain an exclusive meaning, and fourth, the more difficult target words have a suffix which changes the syntactic function of a word but adds no additional substantive cues related to the meaning of the word.

Grade 8: Decoding and Word Meaning

Analysis of Subskills 1, 3, and 4

The average proportion of grade 8 students who correctly answer Decoding and Word Meaning items is displayed in Table 25 for students who reach and do not reach grade level standard, along with the subskill discriminations.

AVERAGE ITEM DIFFICULTY AND ITEM DISCRIMINATION FOR EIGHTH GRADE +R AND -R STUDENTS ON DECODING AND WORD MEANING SUBSKILLS

TABLE 25

Subskill	Average Difficulty for +R Students	Average Difficulty for -R Students	Average Item Discrimination
DW 1	.80	.45	.34
DW 3	.98	.81	.17
DW 4	.85	.51	.34

Contextual Word Meaning (DW 3) is extremely easy for the +R students and it is the subskill on which the -R students have their best performance of all the subskills assessed at grade 8. DW 1 is the most difficult DW subskill for both groups; among all subskills assessed, it is the second most difficult for the +R students. Both the +R and -R students find the DW 4 items slightly easier than the DW 1 items but the DW 4 items are equally discriminating.

Sight Recognition (DW 1): Grade 8

The +R and the -R students have a relatively low performance on DW 1 items. Easier items tend to focus on vocabulary that likely appear more frequently in students' instructional material, while more difficult DW 1 items focus on grade level words that likely appear infrequently. The frequency of appearance of these words likely extends to their appearance in oral language usage as well.

Contextual Word Meaning (DW 3): Grade 8

The +R students perform extremely well in DW 3 items at grade 8, and the -R students exhibit their highest performance in DW 3 among all BSAP Reading subskills. Item analyses reveal no particular pattern with respect to factors affecting difficulty, which is not surprising since the performance within the +R and -R groups on in DW 3 items is quite homogeneous.

Structural Word Meaning (DW 4): Grade 8

The +R students perform well on DW 4 items at grade 8, while -R students have some difficulty with these items. The general familiarity of target root words affects item difficulty; easier items focus on more familiar and more frequently occuring vocabulary words. The target words of easier items possess common prefixes and prefixes which have an exclusive meaning. Such prefixes tend to provide direct cueing of correct answers. The target words of more difficult DW 4 items focus on vocabulary words that, while at or below grade level, have root word forms that are unfamiliar to students, possess prefixes with variable meaning, or use suffixes

which change the syntactic function of target words but add few additional meaning cues.

Summary

Decoding and Word Meaning: Grades 6, 8

DW 1: Sight Recognition

DW 3: Contextual Word Meaning

DW 4: Structural Word Meaning

At grade 6, DW 1 and DW 3 items appear relatively easy for both +R and -R students; both groups, however, demonstrate much lower performance on DW 4 items.

Easier DW 1 items contain target vocabulary words that likely appear more frequently in reading materials, and are therefore more familiar than the words that appear in the harder DW 1 items. Easier DW 3 items tend to be heavily cued by single words in context. More difficult DW 3 items cue the meaning of the unfamiliar target word with a phrase, thus increasing their reading and comprehension load. These more difficult DW 3 items also employ negated phrase cueing where the item asks for the unnegated target word meaning; on these items, students must reason through the task using their understanding of opposition in meaning. Easier DW 4 items tend to have target words that possess more familiar root word forms, while more difficult items possess more unfamiliar root word forms. More difficult DW 4 items also tend to possess target words with prefixes that are divergent (e.g., "con") in their meaning than the easier DW 4 items which contain words with more restricted prefixes (e.g., "pre").

At grade 8 the DW 1 and DW 4 items are relatively difficult for both the +R and -R students, while DW 3 items are relatively easy for both groups.

DW 1 is the most difficult of the Decoding and Word Meaning subskills for grade 8 students. Easier items tend to focus on vocabulary that likely appear more frequently in reading materials, while more difficult DW 1 items focus on words that appear less frequently and are, therefore, less familiar.

DW 3 items at grade 8 are highly homogeneous, revealing little pattern in students' performance. However, easier DW 4 items tend to focus on more familiar vocabulary and vocabulary possessing prefixes with highly restricted meanings. The more difficult DW 4 items focus on less familiar root forms and possess prefixes with variable and divergent meanings. Difficult DW 4 items also can focus on derived word forms using suffixes that change syntactic and grammatic function of the target words, but add little additional direct semantic information.

DETAILS

GRADES 6, 8

Skill Overview

For Details, students read a stimulus selection then must locate and extract explicitly stated information. Detail items are designed to solicit explicit information about characters, events, quantities, frequencies, procedures, and locations. Additionally, Detail items may assess students' ability to determine the correct chronological order for a group of narrated events, including the ability to identify the order of events relative to each other.

Grade 6: Details

Analysis of Skill

At grade 6, the items measuring Details are very easy for +R students (difficulty of .91) and moderately difficult for the -R students (difficulty of .59). The Detail items have a discrimination index of .32.

The relative easy Detail items basically require students to identify a specific piece of information from the stimulus material. The item stems for such questions provide a nearly verbatim repetition of a sentence from the stimulus with the substitution of a wh- word for the piece of information being requested. Other key words from the stimulus are repeated in the stem and constitute very explicit cueing enabling students to match the stem

to a single sentence in the stimulus. Once the stem is matched to a sentence in the stimulus, the correct answer is identified as the word in the stimulus that is replaced in the stem by a wh- word. The relatively easy items have options that contain an average of 2 1/2 words which reflects the fact that the questions ask for specific and isolated details.

For example, imagine a stimulus paragraph of 8-10 sentences which contains the following sentence: "Bill was in Detroit when he got sick." A relatively easy DE item would ask, "Where was Bill when he got sick?" The target word "Detroit" is replaced by "where" and all other key words are repeated. It is quite simple for students to identify the relevant item sentence in the stimulus material by simply matching key words from the stem to the stimulus sentences. The target words in these relatively easy DE items are verbs, nouns or adjectives.

A second type of relatively easy DE item also contains considerable repetition of a key word or words and deals with activities in an explicitly stated sequence of events. The activities in the sequence are explicitly marked by signal words such as "First," "After," "Last," etc. These items contain such extensive and explicit marking of the sequence that students do not have to understand the sequence of events to answer the questions. For example, imagine a stimulus paragraph of 8-10 sentences that contained the sentence, "The first thing Bill did when he got to Detroit was call his friend Mike." Questions like "What was the first thing Bill did when he got to Detroit?" (e.g. "Call his friend Mike") can be answered by the key-word matching strategy previously described.

The more difficult DE items generally do not repeat signal or key words from the selection in the item stem. For example, the question, "What was the first thing Bill did when he got to Detroit?" (e.g., "Call his friend

Mike") can be answered using the stimulus sentence "After he arrived in Detroit, Bill telephoned Mike whom he had known for years." But students must actually understand the sequence of events to answer this question because the sequence signal word "first" in the question does not match the sequence signal word "after" in the relevant stimulus sentence. Students who try to answer this question by matching the word "first" from the stem to the word "first" in the stimulus will be unsuccessful and may actually choose an incorrect answer if the word "first" is used in some other event in the stimulus.

The most difficult DE items also require students to understand stimulus material which describes a series of events. These very difficult items generally ask students to identify an event that occured "just before" or "just after" some other event. With these questions, students must first locate the event used as the referencing event. Second, students must locate the event "just before" or "just after" the referenced event. Third, students must find the event they have identified as the correct answer among the options. This type of item is shown in the example below:

Aunt Mary decided to visit Europe. She went to England for two weeks. Then she crossed the English Channel. She traveled to France, visiting museums along the way. Her next trip was to Germany. She stayed there at an old inn near the city of Berlin.

What did Aunt Mary do just after she visited museums?

(A) went to Europe

**(B) stayed at an inn*

(C) crossed the English Channel

(D) traveled to France

First, students might answer this item incorrectly if they did not correctly locate the reference event "visited museums." Second, students might answer incorrectly if they could not locate the correct event after (or before) the referenced event. And third, students who identify the event "just after Aunt Mary visited museums" as "Aunt Mary traveled to Germany," may answer incorrectly because this apparently correct answer is not among the option choices. In this case students must understand the implied meaning of the question is, "From among the choices listed below, what did Aunt Mary do just after she visited the museums?"

Grade 8: Details

Analysis of Skill

At grade 8, the items measuring details are easy for the +R students (difficulty = .88) and moderately difficult for the -R students (difficulty = .58). The detail items have a discrimination of .30.

The relatively easy grade 8 Detail items have one or more of the following features. First, the item uses the substitution of a wh- word for a key noun or verb asking for the identification of a specific detail. All other words are repeated from the target sentence. Second, the item asks for an event in a sequence of events and repeats signal words like "first," "last," or "before." Third, the item asks for an event "after" some other event. With these easy items, the order of the events in the stimulus material is identical to the order in the stem and answer.

The difficult grade 8 Detail items have no cueing of any sort to help students find the correct answer to the questions. Temporal sequence signaled by words and verbatim repetition of key story elements are absent in the more difficult items.

The difficult items tend to focus on target sentences that contain two events or elements, only one of which is relevant to the question. The additional element in a single sentence disrupts the simple one-event-for-one-sentence structure.

Finally, the more difficult items tend to ask about events "just before" other events. These items require students to search backwards in the event sequence. These items are much more difficult than "just after" questions in which students search for events in the same order in which the events are described in the stimulus material.

Summary

Details: Grades 6, 8

At grade 6, the items measuring Details are very easy for the +R students and moderately difficult for the -R students.

Easier DE items at grade 6 ask for a specific piece of information or an event contained in a sequence of events. The correct answers to these items typically repeat verbatim the necessary information from the stimulus passage, and in the instance of events in a sequence, repeat a signal word from the passage.

More difficult DE items at grade 6 possess little of the key and signal word repetition that characterize the easier DE items. Very difficult DE items tend to ask students to identify an event in a sequence that occurred "just before" or "just after" some other event; items of this type possess no cueing through vocabulary word repetition.

At grade 8, items measuring Details are easy for the +R students, and moderately difficult for the -R students.

Easier DE items at grade 8 tend to possess cueing between item and stimulus passage in the form of vocabulary repetition. More difficult DE items possess little or no such cueing and they focus on target sentences that contain two events or elements, only one of which is relevant to the question posed in the item. Finally, DE items that ask for events "just before" other events are exceedingly difficult.

MAIN IDEA

GRADES 6, 8

Subskill Overview

For Main Idea, the student is presented with a reading selection from which the student must extract the central meaning. At grades 6 and 8, the student must identify two types of main ideas:

1. Paraphrased Main Idea (MI 2). The student must identify the main idea which is a paraphrase of the main idea of the reading selection. The main idea is explicitly stated somewhere in the selection. The degree of paraphrase increases as the grade level increases.
2. Inferred Main Idea (MI 3). The student must identify the main idea which is not explicitly stated anywhere in the selection. The main idea is implied in one or more parts of the selection. The degree of inference increases as the grade level increases.

Both Paraphrased and Inferred Main Idea items require students to perform a number of associated tasks. First, students must identify the major or salient points and details in a reading passage. Second, students must determine whether a "potential main idea statement" is accurate according to passage information. Third, the student must determine whether a "potential main idea statement" is general enough to reflect all of the passage's major or salient points, without being so general that it is excessively broad.

Grade 6: Main Idea

Analysis of Subskills 2 and 3

The average proportion of grade 6 students who correctly answer Main Idea items is displayed in Table 26 for students who reach and do not reach grade level standard, along with the subskill discriminations.

AVERAGE ITEM DIFFICULTY AND ITEM DISCRIMINATION FOR SIXTH GRADE +R and -R STUDENTS ON MAIN IDEA SUBSKILLS

TABLE 26

Subskill	Average Difficulty for +R Students	Average Difficulty for -R Students	Average Item Discrimination
MI 2	.74	.39	.35
MI 3	.75	.37	.38

The MI subskills at grade 6 are among the most difficult BSAP subskills for both +R and the -R students, and are high in their ability to differentiate between the two groups.

Paraphrased Main Idea (MI 2): Grade 6

The relatively easy MI 2 items at grade 6 are associated with reading selections that have one or more of the following characteristics. First, the salient information or key sentence needed to identify the main idea appears early in the selection and is contained in a single introductory

topic sentence. Second, the key topic sentence cues the correct option through repetition of key vocabulary words and phrases; therefore, the paraphrasing required of students is minimal and usually of a simple synonym substitution type. Third, the remainder of the sentences in the selection provide a high degree of support for the topic sentence and correct option, as well as occasional repetition of the salient main idea information.

The type of Paraphrased Main Idea transformation that students must make in the easier MI 2 items can be illustrated by examining the hypothetical topic sentence in the box below.

Example 1

> Some researchers have invented a new machine that records a person's brain activity. [Several sentences elaborating upon this topic sentence follow.]

What is the main idea of this passage?

ans: A new device helps researchers monitor people's brain activity.

In this example, students must identify the paraphrased main idea. The correct answer differs from the passage topic sentence in two ways. The syntax of the sentence construction is slightly altered and "monitor" is substituted for "records." Students have the least difficulty with this type of MI 2 item.

The relatively more difficult MI 2 items at grade 6 accompany reading

selections that have one or more of the following features. First, these selections often consist of two paragraphs that are implicitly related to one another (see example on page 136). Typically, the first paragraph is very general and does not provide sufficient information to infer the paraphrased main idea. The second is more specific and contains the salient information needed to identify the paraphrased main idea. Another type of passage construction for the more difficult MI 2 items requires students to combine salient information from each paragraph or from beginning and ending positions in one paragraph. While more difficult MI 2 items may have different passage constructions, they all place heavier inter-sentence and inter-paragraph reading demands on the student. These demands require higher levels of inference.

Second, the harder MI 2 items often have reading selections in which some of the salient information needed to identify the paraphrased main idea is presented in figurative or idiomatic language. In addition, the correct option in some of the MI 2 items presents the paraphrased main idea in figurative or idiomatic language. In both situations, students unfamiliar with the idiom will have much more difficulty with these MI 2 items since comprehending the particular idiom involved is necessary to correctly identify the paraphrased main idea.

Third, more difficult MI 2 items have reading selections that consist entirely of descriptive information and do not follow a topic - sentence/supporting-detail sentence construction. The correct answer for such items often repeats one or two salient details from the selection and, although correct, appears to have too narrow a focus. Students seem to have difficulty identifying the most salient points from the selection when this type of passage construction is used.

Fourth, the more difficult MI 2 items typically have one, and sometimes two, incorrect options that are chosen by a high percentage of the students, particularly by the -R students. These popular wrong answers are of several types. First, in the two paragraph construction where the first paragraph is very general and does not possess the salient information necessary to identifying the paraphrased main idea, the most popular wrong answers tend to be inferred main idea statements that are related to the more general paragraph. These wrong answers incorporate some of the salient passage information but not enough to adequately communicate the main idea of the selection.

A second type of popular wrong answer repeats a detail that was explicitly presented in the selection. As the difficulty level of the MI 2 task increases, it appears that students who are unable to perform the task tend to choose wrong answers that they can verify as true details whether or not those answers express the main idea of the selection. Some very difficult MI 2 items possess incorrect options that are substantively similar to the correct answer, but slightly more narrow in their focus; these wrong answers draw a high percentage of student responses. A final type of popular wrong answer consists of an incorrect option that expresses a reasonable inference based on passage information, but is only slightly related to the main idea or far too general to be an adequate main idea statement.

Below is an illustration of the stimulus passage construction described earlier in which the first paragraph is written at a very high level of generality; sufficient salient information for identifying the correct main idea option appears in paragraph two. The reader should note the implicit relationship between these two paragraphs, and that the salient passage

information needed to identify the paraphrased main idea is located in the first sentence of the second paragraph.

Example 2

> Modular building refers to the method of constructing pieces of houses, or modules, at a central location. The housing shortage in this country is a serious problem. Some home designers and builders think that modular building can help.
>
> A factory that will produce modular homes has been started in Calhoun County. It will be constructed near Interstate 95. According to the company's president, this location is ideal. This will allow the company easy access to the interstate highway system so that they can easily ship house modules by truck. A fabrication plant will also be built which will make use of the abundant sand that is available at the site.

What is the main idea of this selection?

KEYED OPTION: (A) Calhoun County will be the location of a new modular home factory.

VALID INFERENCE: (B) Home designers and builders think that more modular homes will be built in the future.

PARAGRAPH ONE MAIN IDEA: (C) New building practices will be required to help deal with the critical need for new housing.

TOO NARROW IN FOCUS: (D) A new fabrication plant will be built in Calhoun County.

The example above is also designed to illustrate the nature of the popular wrong answers that children tend to choose. For example, Option B is a perfectly valid inference based on passage information, but does not communicate adequately the main idea. Option C could be construed to be the

inferred main idea of paragraph one, but is too general to be a proper main idea statement encompassing both paragraphs. Option D is substantively similar to the correct answer, but is too narrow in its focus since it mentions only a subordinate element; that is, the fabrication plant is subordinate to the modular home factory, since it is only one element of the factory.

Inferred Main Idea (MI 3): Grade 6

The relatively easy MI 3 items at grade 6 have reading selections that have several characteristic paragraph structures. The first consists of a passage that provides an eleborated concrete description without a general level topic sentence. This type of passage provides a wealth of specific concrete support for the correct Inferred Main Idea, and has minimal inter-sentence and inter-paragraph reading demands. The correct option in such items differs from incorrect options with respect to level of generality.

A second type of passage construction found with easier MI 3 items repeats a salient piece of information three or more times in the selection. This repetition appears to provide a signal or cue to the student that is helpful both in identifying the correct option and in eliminating incorrect options.

A third type of passage construction used with the relatively easier MI 3 items describes a sequence of events or a factual account of a situation. The correct answer for such items tends to be a summary statement that renames the salient event or fact from the selection. Popular incorrect options for these items present fairly reasonable but overgeneralized inferences based on passage content.

The more difficult MI 3 items have certain features which distinguish them from easier MI 3 items, although all MI 3 items are generally quite difficult for sixth graders. The single most striking characteristic of the very difficult MI 3 items is the type of reading selection they contain. These reading selections typically consist of two paragraphs that are linked causally. A first paragraph may present either the "cause" or the "effect" element of the causal link, while the second paragraph presents the other member of the causal link as well as supporting details. This particular passage construction dominates the reading selections in MI 3 at grade 6, and is consistently associated with very difficult MI 3 items. This is probably due in large part to the fact that this construction places both heavier inter-sentence and heavier inter-paragraph reading demands on students (see example on the next page).

The more difficult MI 3 items typically have one and sometimes two incorrect options that are chosen by a high percentage of students in general, and a very high percentage of the -R students in particular. These popular wrong answers are of several types. First, incorrect options that repeat numerical information from the selection frequently are chosen. Such numerical information is visually salient and, for students having difficulty with the skill, represents passage information that they can identify as true using a simple visual matching strategy. Second, popular incorrect answers tend to be associated with the last sentence of the stimulus passage regardless of the level of generality or information content. Students appear to be overly reliant on positional cues or overly influenced by recency effects in getting the inferred main idea.

The example below is included to illustrate the inter-paragraph causal link construction previously described. The accompanying incorrect options illustrate the types of popular wrong answers.

Example 3

> For many years, houses have been built by teams of men right on the land where the houses would remain. Now many house builders are doing their jobs in a completely different way. This different way of building houses is not at all new, however.
>
> In days gone by, the unfortunate news was high costs. Shipping building materials from many suppliers to many different locations where houses are to be constructed is getting too expensive for builders and buyers alike. It's much less costly to build homes in one central place, then ship them. The builder first orders supplies to make perhaps 50 homes. Then the builder constructs the homes in modular style, so that each house is a collection of several modules or pieces. The modules can be shipped by truck to the building site. At first, there were problems with this method, but they have now been solved. Builders today can construct an average sized modular home for about $12,000 less than a house built by the old method.

What is the main idea of this selection?

KEYED OPTION: (A) Builders are building more modular homes because they are less expensive than building on location.

CORRECT LAST SENTENCE DETAIL: (B) Modular homes cost $12,000 less than homes built on-site.

PLAUSIBLE BUT OVER-GENERALIZED INFERENCE: (C) Home builders and home buyers like modular homes better than old fashioned homes.

CONTRADICTORY: (D) The cheapest way to build a new home is to construct it on-site.

The reader should note that in this sample passage, the two paragraphs taken separately have two different main ideas. Also, the "new way" of house building mentioned in paragraph one is not explicity described in paragraph one but is explicated in paragraph two. Paragraph two probably provides sufficient information to infer the main idea. However, paragraph one serves as an organizer for paragraph two by implicitly aiding the reader in sorting out the information in the second paragraph. Not all readers are aided in this way by paragraph one, however.

The type of incorrect option illustrated in Option B tends to be a popular wrong answer since it presents a passage detail and the detail is both a piece of numerical information and is drawn from the last sentence in the passage. The type of wrong answer illustrated in Option C tends to be a popular option type since it is a somewhat reasonable inference based on passage information and is like a main idea statement in terms of its level of generality. The contradictory option type demonstrated in Option D tends not to draw a very high percentage of student response.

There are both MI 2 and MI 3 items which use two paragraphs in their reading selections. The first paragraph relates some general observations about the topic while the second is far more specific. The relationship between the two paragraphs is not marked by signal words nor explicitly stated. The second paragraph could easily be subordinate to the first and this relationship could be clearly marked if the second paragraph began, "For example, ..." In such a case, the main idea would have to come from the first paragraph.

The lack of a clear relationship between paragraphs undoubtedly makes

it difficult for students to order ideas as superordinate main ideas or as subordinate supporting ideas. Therefore, part of the students' task is "meta-structural" in that students must order the information from the paragraphs as superordinate or subordinate without cues from signal words. Students may be looking for main ideas in paragraphs that are general in nature because they have the basic concept that main ideas are general statements.

Grade 8: Main Idea

Analysis of Subskills 2 and 3

The average proportion of grade 8 students who correctly answer Main Idea items is displayed in Table 27 for students who reach and do not reach grade level standard, along with the subskill discriminations.

AVERAGE ITEM DIFFICULTY AND ITEM DISCRIMINATION FOR EIGHTH GRADE +R AND -R STUDENTS ON MAIN IDEA SUBSKILLS

TABLE 27

Subskill	Average Difficulty for +R Students	Average Difficulty for -R Students	Average Item Discrimination
MI 2	.83	.52	.31
MI 3	.82	.47	.35

The MI subskills at grade 8 are moderately easy for the +R and difficult for the -R students, and differentiate the two groups quite well.

Paraphrased Main Idea (MI 2): Grade 8

The difficulty of MI 2 items at grade 8 is related to the topic or content of the reading selections and structure of the reading passages. Easier MI 2 items have familiar passage topics like a common type of animal, and elaborate upon and describe unfamiliar features of the more familiar topic. More difficult MI 2 items have passage topics and content that appear to be far less familiar. More difficult MI 2 items also have familiar passage content that is presented in an unfamiliar context. For example, such a passage might present relatively obscure information about a famous person who is very well know for something entirely different. In these instances, students appear to rely heavily on their prior knowledge to identify the paraphrased main idea.

Relatively more difficult MI 2 items have reading passages that present a series of events in mixed order relative to the order in which the events would have occurred naturally. Students' low performance of this sort of item may occur for several reasons. The mixed ordering on events in a sequence makes identifying the paraphrased main idea a more difficult task. The other is that paraphrased main ideas for a sequence of events passage require students to identify the salient or "most important" event(s) from a series of 7 to ten events, and that this is a more difficult kind of operation. Additionally, relatively more difficult MI 2 items have reading selections which tend to contain several unfamiliar but grade level appropriate vocabulary words that are essential to identifying the paraphrased main idea. Also, more difficult MI 2 items very seldom repeat key words and signal words in the correct answer that appear in the stimulus passage.

In addition to vocabulary cueing, easier MI 2 items tend to contain explicit cueing of topic sentences. This topic sentence cueing is of two types. First, topic sentences can be cued by question sentences embedded in the passage; for example, in the following paragraph the second "question" sentence directly cues the sentence that precedes it as the topic sentence:

Example 4

QUESTION SENTENCE CUEING:	When animal behavior changes drastically, it can tell us about future events. What are some of these strange behaviors? Erratic movement, interruption of regular feeding and mating habits, and listlessness in many animals clue the coming of earthquakes.

Second, topic sentences can be explicitly cued by signal word phrases:

Example 5

SIGNAL PHRASE CUEING:	Some signals indicating the coming of an earthquake can be observed in a variety of animals. For example, fish begin to swim erratically and jump out of the water.

In the signal word phrase cueing illustrated above, the phrase "for example" in example 5 cues the reader that sentence 2 not only contains factual examples, but also that the prior sentence is at a higher level of generality, and is therefore likely to be the topic sentence. Related

kinds of cueing occur when passages use phrases like, "The scientists discovered that . . ." Students seem to understand that the phrase or sentence that follows "the scientist discovered that . . . " is likely to be a topic sentence. Students appear to use these phrases to help locate both topic sentences and salient information that is likely to relate to the main idea.

Relatively easier MI 2 items tend to have reading passages that present the salient topic sentence in the first paragraph, or within the first three sentences in one-paragraph reading selections. These easier MI 2 items typically require simple paraphrasing that involves straight-forward, word-for-word synonym substitutions. In contrast to these easier items, the more difficult MI 2 items require students to paraphrase clauses and phrases and to superordinate multiple topic sentences from the different paragraphs. Generally, the reading selections of the easier MI 2 items provide a large degree of support for the main idea.

Item difficulty is affected in large part by the nature of the incorrect option choices. Two types of incorrect options tend to be chosen by a high percentage of students. First, wrong answers that repeat a key vocabulary word from the selection are chosen frequently. Second, students frequently choose wrong answers that are either valid inferences or overstate reasonable generalizations based on passage content. A high percentage of the -R students choose these kinds of wrong answers. These types of wrong answers suggest that the -R students are relying on a "match-the-word" strategy which leads to choosing options that repeat key vocabulary and that -R students have a partial understanding that main idea statements are somewhat general types of statements.

Inferred Main Idea (MI 3): Grade 8

The most striking feature of the relatively easy MI 3 items at grade 8 is what they are missing, not what they contain. In general, these relatively easy items do not require combining inferences from each paragraph in the stimulus. Sometimes the main idea can be inferred from only one of the paragraphs in the selection. When information from each paragraph is needed, the information tends to be fairly explicit in the paragraphs and requires little or no inference. Such items require inter-paragraph information combining but do not require inference skill.

Also absent from the easy MI 3 items is an incorrect answer which repeats a salient detail from the stimulus passage. The options are homogeneous in that they tend to be general statements of the type that could be a main idea. No one incorrect answer stands out as particularly attractive by repeating a key detail from the passage.

The most difficult MI 3 items at grade 8 have three common characteristics. First, these Inferred Main Idea items require students to combine information from each of the paragraphs in the stimulus material. Second, the information needed from the several paragraphs is not explicitly stated but requires some degree of inference. Third, the more difficult MI 3 items have one or more incorrect answers which are appealing to all students but are especially appealing to -R students. These attractive incorrect answers are either the main idea for a single paragraph or a single salient detail repeated from the passage.

Summary

Main Idea: Grades 6, 8

MI 2: Paraphrased Main Idea

MI 3: Inferred Main Idea

The MI subskills at grade 6 are among the most difficult in BSAP for both +R and -R students at this grade level.

Easier MI 2 items tend to have clear and easily identifiable topic sentences that are only slightly different in wording from the keyed option. More difficult MI 2 items tend to have one or more of the following charateristics: a two paragraph stimulus passage with the second paragraph containing the necessary information for identifying the paraphrased main idea; salient main idea information expressed in the context of idiomatic language; a stimulus passage that requires students to identify salient information/events because there is no topic sentence; and/or highly popular incorrect options.

The easier MI 3 items and reading selections have several common features. The stimulus selections for these relatively easy items may possess no topic sentence, but the passage provides a great deal of concrete support for the main idea, often through repetition of salient information. Also, easier MI 3 items tend to have correct answers which basically summarize the selection. More difficult MI 3 items have reading selections with two paragraphs that are causally linked. Identifying the Inferred Main Idea for these reading selections requires students to com-

bine salient information from each paragraph, thus increasing the inter-paragraph comprehension demands of these passages. Difficult MI 3 items also tend to have very popular wrong answers.

The MI subskills at grade 8 are moderately easy, relatively speaking, for the +R and difficult for the -R students.

Easier MI 2 items at grade 8 have several common features. The reading passage in these items tends to be highly familiar, and correct answers tend to be word-for-word paraphrases of topic sentences from the passage. Also, easier MI 2 items have reading selections in which the location of the topic sentence is explicitly cued or marked.

More difficult MI 2 items have one or more of the following features. Reading selections present a sequence of events in mixed order and possess far less cueing of topic sentences; reading passages often contain vocabulary words that are unfamiliar but important in identifying the main idea; correct answers often paraphrase clauses or phrases from the reading selection in contrast to word-for-word paraphrasing; and often the MI 2 task requires the student to superordinate multiple topic sentences.

Easier MI 3 items at grade 8 tend to require students either to make inferences from just one paragraph or short segment of the stimulus reading passage, or to combine literal information from across paragraphs. More difficult MI 3 items tend to require a greater degree of inference, the combining of inferred information from across paragraphs of the reading passage, and possess one or more incorrect options that are very appealing to students.

REFERENCE USAGE

GRADES 6, 8

Subskill Overview

For Reference Usage, the student must select or use a variety of reference sources. As grade level increases, eligible reference sources expand to include more sophisticated sources. Additionally, the level of detail of the information requested from each source becomes increasingly complex as grade level increases.

1. Selecting a Reference Source (RE 1). The student is presented an information need and must select the information source most likely to contain the information requested. At grade 6 eligible reference sources include: table of contents, dictionary, calendar, map, chart, encyclopedia, card catalog, book index, telephone directory, newspaper classified ads, glossary, graph, and table. Grade 8 sources expand to include: newspaper index, bibliography, appendix, and readers' guide.

2. Using a Reference Source (RE 2). The student is presented an excerpt from a reference source and must extract the requested information. Eligible reference sources include those cited in the description of RE 1.

Grade 6: Reference Usage

Analysis of Subskills 1 and 2

The average proportion of grade 6 students who correctly answer Reference Usage items is displayed in Table 28 for students who

reach and do not reach grade level standard, along with the subskill discriminations.

AVERAGE ITEM DIFFICULTY AND ITEM DISCRIMINATION FOR SIXTH GRADE +R and -R STUDENTS ON REFERENCE USAGE SUBSKILLS

TABLE 28

Subskill	Average Difficulty for +R Students	Average Difficulty for -R Students	Average Item Discrimination
RE 1	.90	.61	.29
RE 3	.93	.70	.23

The RE subskills at grade 6 are among the easiest in BSAP for both the +R and -R students. RE differentiates moderately well between the two groups.

Selecting A Reference Source (RE 1): Grade 6

The relatively easy RE 1 items deal with telephone directories, dictionaries or encyclopedias. In all cases, students are likely to be familiar with the particular reference source required by the item. The nature of the tasks tends to make it easy to exclude a number of wrong answers. For example, questions of the form, "where would you look for a picture of ...", are clearly restricted to a reference source with pictures. Thus, reference sources unlikely to contain pictures can be easily eliminated.

The more difficult RE 1 items deal with book indexes, call numbers for a book (card catalogs) and classified ads. Students tend to confuse a book's index and table of contents with respect to the information they contain.

Students do not know at grade 6 that only a book's index is an alphabetical listing. When asked where they would look for the "call number" of a book, -R students commonly select a "telephone directory." Their choice of incorrect answer would suggest that students are associating "call" from "call number" with "call" from "calling a telephone number." "Telephone directory" also is a popular wrong answer to the question, "where would you look for a job opportunity?" This is a plausible answer in the experience of some children.

Telephone directories may be the reference source with which students are most familiar and have the most direct experience. Questions in which "telephone directory" is a correct answer are likely easier as a result of this familiarity; questions in which "telephone directory" is a wrong answer may be especially difficult because students overgeneralize the usages of telephone directories since other reference sources are less familiar. Students' difficulty with questions about parts of a book and card catalogs may reflect a general lack of school and home experience with these particular reference sources.

Using a Reference Source (RE 2): Grade 6

The relatively easy Reference Usage items at grade 6 can be answered with very little knowledge of the reference source or its proper use. These easy items can be answered correctly by matching key words or phrases in

the item to the same words or phrases in the stimulus material. After the correct match is made, the student then reads the volume number, column heading, row label, adjacent telephone number, or address that goes along with the key word.

A simplified example of this type of item follows.

Example 1

COLOR OF FRUITS ON PLANET X

FRUIT	COLOR
Nurg	Red
Spla	Orange
Noti	Green
Zaplo	Blue

If you were eating a blue fruit on planet X, what would you be eating?

(A) Nurg

(B) Spla

(C) Noti

**(D) Zaplo*

These easy items deal with encyclopedias, charts, telephone directories and classified ads. Because of the nature of this skill many of these items can be answered using a general "point-and-move-your finger-then-read-the-label" strategy.

The more difficult RE 2 items cannot be answered using the simple visual matching strategy just described. Among the difficult items are

questions in which students must identify the proper page number in an index for a particular entry. The index entries are listed with several levels of subdivisions. Students do not seem to be able to follow the superordinate/subordinate relationships among elements in the key words and phrases. For example,

Example 2

On what page would you find information about the history of bridge construction?

	Page	
Bridges----------	79	
Construction----------	82	
Blue Prints----------	83	
History----------	87	*Key
Pictures----------	109	
Demolition----------	57	
History----------	62	
Pictures----------	71	
History----------	126	

Popular wrong answers to such a question would be page 79 (Bridges), page 82 (Construction), and page 126 (History).

The simple visual matching strategy also fails students when they must identify the volume number of an encyclopedia for an entry that does not begin with one of the guide-letters from the set. For example, if Volume 7 had the guide letter R - U, students would have difficulty identifying it

as the key to the question, "In which volume would you find information about Adlai Stevenson?" Students would find far easier the question, "In which volume would you find information about Eleanor Roosevelt?" since the letter "R" is repeated between the guide letters in the stimulus and the surname in the stem.

GRADE 8: REFERENCE USAGE

Analysis of Subskills 1 and 2

The average proportion of grade 8 students who correctly answer Reference Usage items is displayed in Table 29 for students who reach and do not reach grade level standard, along with the subskill discriminations.

AVERAGE ITEM DIFFICULTY AND ITEM DISCRIMINATION FOR EIGHTH GRADE +R and -R STUDENTS ON REFERENCE USAGE SUBSKILLS

TABLE 29

Subskill	Average Difficulty for +R Students	Average Difficulty for -R Students	Average Item Discrimination
RE 1	.90	.59	.31
RE 2	.92	.64	.28

The RE subskills at grade 8 are among the easiest in BSAP for both +R and -R students. RE differentiates moderately well between the two groups.

Selecting Reference Sources (RE 1): Grade 8

Variation in the difficulty of RE 1 items is not simply associated with particular reference sources, but also is associated with the nature and content of the questions. In general, relatively easy items deal with the use of a card catalog, telephone directory, or dictionary. The more difficult items deal with book bibliographies, newspaper indexes, or classified ads.

It is informative to examine the easy items more carefully. The easy card catalog items have parts of books for all the wrong answers and these wrong answers do not attract many students. The easy telephone directory questions ask about the reference source needed if you wanted to call someone. This use of a phone directory is so familiar that it represents an item quite different from all others. The easy dictionary questions asked students to identify the reference sources needed to divide words into syllables. The relative low difficulty of this item may be due to the fact that no other reference source could be a reasonably plausible answer to this question.

In contrast to these relatively easy items, the more difficult RE 1 items seem to have many plausible answers among the options. For example, consider the question, "What tells you where movie information is located in a newspaper?" The correct answer is "newspaper index" and a popular incorrect answer is "a chart." This is a difficult question because most newspapers present their index in chart form and, hence, "chart" is not clearly a wrong answer. Somewhat similar is an item that might ask, "Where can you find a listing of cars for sale?" The correct answer is the "classified ads," but a popular incorrect answer is a "newspaper index."

Students who select "newspaper index" are apparently reading the question as, "Where can you find (out the location of) a listing of cars for sale?" The answer to the question interpreted this way is "newspaper index." A final difficult item type in RE 1 deals with the information contained in a book bibliography. The incorrect answers include "table of contents," "index," and "encyclopedia." Students' responses to this item type suggest that they have rarely used a book's bibliography and are very unfamiliar with its function.

Using A Reference Source (RE 2): Grade 8

Students can answer the relatively easy RE 2 items by matching key words or phrases from the stimulus materials to the same word or phrase repeated in the stem. The reference sources for these relatively easy items include a bibliography, card catalog, a chart and classified ads. These items can be answered using the "match-the-words-and-move-your-finger-to-the-label" strategy. A comprehensive knowledge of the reference source and its uses is not essential.

The more difficult RE 2 items involve the use of a book index, map, dictionary, classified ad, and newspaper index. Book index questions have three levels of subordination and students select answers that indicate they are not using all the levels of information. Map items asks students to identify a location based on two directions, for example, south west of a starting point. Students often pick the option which is south but not west. Dictionary items require students to use the pronounciation key at the bottom of the page, a requirement largely ignored by students. Classified ad questions have correct answers which have three levels of

specification; e.g., "private swimming lessons." Students need to use all levels of information to answer these questions and they generally choose wrong answers which indicate they are using one or two levels only.

Although these difficult items involve many different reference sources, they have a clear common element. All of these items involve the simultaneous use of two or more pieces of information. Students in general, and the -R students in particular, are not able to combine these sources of information to find the keyed option.

A final difficult type of item involves the use of a newspaper index and asks, for example, "Where would you find a writer's opinion about the BSAP program?" The correct page number in the index is next to "Editorial." Students' difficulty with this item almost surely derives from the fact that they are unfamiliar with the word "editorial."

Summary

Reference Usage: Grades 6, 8

RE 1: Selecting a Reference Source

RE 2: Using a Reference Source

Relatively speaking, RE subskills are among the easiest for both +R and -R students at grade 6. In RE 1, students appear to have a firm understanding of the types of information contained in telephone directories, dictionaries, and encyclopedias. On the other hand, they exhibit confusion with the types of information contained in book indexes, card

catalogs, and classified ads; this is particularly true for the -R students.

Relatively easy RE 2 items can be answered by matching a key vocabulary word from the item stem to the same word in the stimulus. RE 2 items that cannot be answered via this matching strategy tends to be more difficult. In particular, RE 2 items that require students to sort through different levels of subordination/superordination as typically exist in a book index tend to be more difficult.

In RE 1 items, grade 8 students appear to be familiar with the type of information contained in telephone directories, card catalogs, and dictionaries. However grade 8 students, particularly -R students, demonstrate confusion with the kinds of information contained in book bibliographies, newspaper indexes, and classified ads. The relative plausibility of wrong answers greatly influences item difficulty in RE 1 items. Students, and particularly -R students, demonstrate some confusion in differentiating between specific reference sources (e.g., a newspaper index) and more general terms that might be considered equivalent (e.g., "chart" since a newspaper index is a type of chart) in RE 2 items.

INFERENCE

GRADES 6, 8

Subskill Overview

For Inference items, students must read the selection presented and then determine the correct inference based on the selection. As grade level increases, the level of inference increases. The reading passages become longer, and the types of passages become more complex. The level of inference required of the student also increases. The four IN subskills are defined as follows:

1. Making Comparisons (IN 1). Students must compare persons or objects relative to some characteristic or quality. The number of objects/persons compared, as well as the number of attributes in question, increases with grade level.

2. Determining Cause and Effect (IN 2). Students must read a passage and then determine the most reasonable or plausible cause of a given effect or situation. As grade level increases, the relationship between the cause and the effect becomes less direct with respect to the description of events and the circumstances surrounding those events.

3. Drawing Conclusions (IN 3). Students must read the selection and then determine the most reasonable or plausible conclusion that can be drawn based on the information in the passage. The conclusion must be supported by the information contained in the passage. The conclusion is a general statement based on a passage that develops a single topic with details.

4. Predicting Outcomes (IN 4). Students must read the passage and then determine the most reasonable or plausible action that could continue or conclude the series of actions described. Difficulty increases across grade level with changes in the kind of passage information given and the level of inference required.

In grade 6, reading passages consist of simple non-fiction selections. In grade 8, reading passages can be comprised of essays, editorials, articles, and letters.

Grade 6: Inference

Analysis of Subskills 1, 2, 3, 4

The average proportion of grade 6 students who correctly answer Inference items is displayed in Table 30 for students who reach and do not reach grade level standard, along with the subskill discriminations.

AVERAGE ITEM DIFFICULTY AND ITEM DISCRIMINATION FOR SIXTH GRADE +R and -R STUDENTS ON INFERENCE SUBSKILLS

TABLE 30

Subskill	Average Difficulty for +R Students	Average Difficulty for -R Students	Average Item Discrimination
IN 1	.81	.40	.41
IN 2	.82	.40	.42
IN 3	.73	.27	.46
IN 4	.81	.36	.45

The inference items are difficult for both the +R and -R students. All four subskills are sharply discriminating. The Drawing Conclusions (IN 3) items are especially difficulty and discriminating. The -R students'

performance on IN 3 is barely above the level of performance expected by students randomly selecting answers (.25).

Making Comparisons (IN 1): Grade 6

The relatively easier IN 1 items and their accompanying stimulus selections at grade 6 differ from more difficult IN 1 items and passages in several ways. The stimulus passages for the easier items tend to explicitly communicate the basis or dimension of comparison (e.g., tallest; quickest). People, objects, or procedures described in the passages are given specific characteristics. When these characteristics are relevant and concrete, the dimension of comparison tends to be explicit and the accompanying items easier. Often, these easier items ask students for comparisons that involve inferring the "best procedure" or method for doing something. The more difficult IN 1 items tend to be accompanied by passages that imply but do not actually name the dimension of comparison.

The degree to which passages are restricted affects overall passage clarity and influences item difficulty. Easier items tend to have more restricted or convergent reading passages. These passages limit the nature of the comparisions that one might be asked to make based on the passage. Unrestrictied or divergent passages, in contrast, can be used as the basis for several different types of comparisions. In the items with divergent passages, the dimension of comparison is often not clear until the item stem is read. In addition, in these more difficult IN 1 items, relevant comparison information in the passage is often paraphrased in the item in two and three word phrases ("moves slowly" is equivalent to "loses momentum") that must be translated onto the dimension of comparison (Which

is the fastest?).

The easier IN 1 items tend to have incorrect options that are explicitly and clearly negated in the reading passage, and these are often cued in the passage by signal words like "but," "however," etc. Easier items also tend to have correct answers that are explicitly confirmed in the passage.

Nearly all IN 1 items, and particularly the more difficult ones, require students to combine information from several passage sentences to make the comparison. The response patterns of -R students suggest they are not engaging in the inter-sentence combining of information required by IN 1 items.

Determining Cause and Effect (IN 2): Grade 6

Both easy and hard cause and effect (IN 2) items present students an effect in the item stem, and ask students to identify a cause or reason for the effect. The easier IN 2 items differ from the more difficult IN 2 items in several ways. The easier items tend to ask about very concrete or specific effects (e.g., What is the most likely reason the tree fell down?). More difficult IN 2 items ask about more abstract or general effects (e.g., What is the most likely reason for Jim's success?). More difficult IN 2 items which ask about specific effects, often have several plausible reasons or causes, only one of which is the "best" answer.

The more difficult IN 2 items also tend to require more inter-sentence combining of information, contain little or no signal word cueing of salient information, and occasionally miscue salient information with signal words. In the instance of one of the more difficult IN 2 items, the

stem is an inference based on the passage, which asks for an inferred cause (from the passage) for the inferred effect stated in the stem.

The easier IN 2 items occasionally require little more than paraphrasing a salient passage sentence. The easy IN 2 items also possess wrong options that are explicitly negated or explicitly contradicted in the passage.

The wrong answers students choose in the IN 2 items is quite revealing. The -R students appear to be either unclear regarding the task or unable to extract enough meaning from the selection to identify plausible answers. Often, for example, in items that ask for an inferred cause they choose an incorrect option that states an effect, not a cause. The confusion of the -R students in IN 2 items is further suggested by the fact that they tend to choose incorrect answers that are true in their own experience. Both +R and -R students who get IN 2 items wrong tend to choose wrong answers whose content comes from the sentence immediately following the passage's statement of effect. For example,

> Mr. Hall opened his first clothing store in 1923. Within ten years he became the most successful clothing retailer in the country. Mr. Hall learned to sew when he was just four years old. (The passage goes on to say that Mr. Hall produces quality goods at a fair price).

What is the most likely reason for Mr. Hall's success?

The focus of the question is "Mr. Hall's success," which is established in the second sentence. Students answering this question incorrectly would

generally find incorrect options like "He learned to sew at a very early age" extremely attractive. This is particularly true for the -R students. The contiguity of sentences 2 and 3 in the example above apparently suggest a causal connection between the two sentences to many students. Students who pick this wrong answer are not identifying salient passage information or using inter-sentence comprehension strategies.

Drawing Conclusions (IN 3): Grade 6

IN 3 at grade 6 is the most difficult of all BSAP subskills for the -R students, and one of the most difficult subskills for the +R students. There is so little variation in the difficulty of IN 3 items, that comparisons of "easier" to "harder" items reveals little because all the items are "harder." Analysis of IN 3 items, however, does reveal a number of factors that likely contribute to their consistent difficulty.

First, IN 3 items at grade 6 use a common stem throughout:

Which of the following conclusions can most reasonably be drawn from this selection?

The stem itself represents a difficult reading comprehension task. It is both difficult with regard to its vocabulary load and its level of abstractness. Students who do not have experience with this particular question before seeing it on the examination may not understand what the question is asking them to do.

Further, IN 3 items at grade 6 represent extremely difficult reading tasks. All item options tend to be lengthy thereby increasing their

reading load. More critically, many item options are constructed to be superordinate statements that encompass most or all of the passage information. These options, therefore, are usually expressed at a high level of generality, resemble summaries, and are highly dependent on inter-paragraph comprehension. Other items that are more specific and less superordinated also focus on salient passage events. Thus on all items, students must be able to recognize the salient events and description in the passage to correctly identify a "reasonable conclusion." In the most difficult IN 3 items they must recognize the salient events and description in summarized or superordinate form.

Lastly, in the most difficult of the IN 3 items, the passage presents a series of events out of order and the events span several centuries. The correct option for this item is a superordinated summary of the salient passage events and description. This "juggled sequence of events" passage structure appears to further compound the already difficult skill of Drawing Conclusions.

Students' weak performance in IN 3 at grade 6, particularly that of the -R students, suggests that they do not understand the nature of the subskill and/or are unable to demonstrate the skill in the context of these particular items and passages. This is evidenced by their response patterns, in which they tend to choose incorrect options that are true in their own experience.

Predicting Outcomes (IN 4): Grade 6

The most striking difference between easier and more difficult IN 4 items is the structure of the respective reading passages upon which they

are based. Easier IN 4 items have passages which describe in sequence a series of actions a single character takes, and ask students for "what most likely happened next?" The overt actions described in the passage converge on a relatively small number of plausible "next events," thus restricting the range of possible correct answers.

More difficult IN 4 items are associated with passages that generally have a more complex story line, and have one or more of the following specific features. First, these passages often introduce two characters, either of whom could perform the likely next event. Second, these passages typically introduce salient character information and description that becomes central to identifying the most likely next event. Third, descriptive information about a character's motivation or internal state might change, representing a change in attitude of the character that in turn influences the most likely next event. Fourth, the story line developed in such passages is less restricted than in easier passages, allowing for a wider range of events that could occur next. This last factor increases the inter-paragraph reading demands of difficult reading passages.

Difficult IN 4 items at grade 6 contain options that possess certain characteristics. For example, more difficult items occasionally have as the correct answer an event that is not "most likely" in general, but is the most likely next event listed among the options. Often, incorrect options that draw a high proportion of student responses, particularly from among the -R students, are actually correct based on the reading of one stimulus paragraph, but are not correct based on the entire passage.

Grade 8: Inference

Analysis of Subskills 1, 2, 3, 4

The average proportion of grade 8 students who correctly answer Inference items is displayed in Table 31 for students who reach and do not reach grade level standard, along with the subskill discriminations.

AVERAGE ITEM DIFFICULTY AND ITEM DISCRIMINATION FOR EIGHTH GRADE +R and -R STUDENTS ON INFERENCE SUBSKILLS

TABLE 31

Subskill	Average Difficulty for +R Students	Average Difficulty for -R Students	Average Item Discrimination
IN 1	.87	.51	.36
IN 2	.87	.40	.47
IN 3	.78	.33	.45
IN 4	.83	.42	.41

The IN 1 and IN 2 subskills are relatively easy for the +R students and moderately difficult for the -R students. IN 3 and IN 4 are both more difficult for the +R students than IN 1 or IN 2. IN 2 is particularly discriminating (.47) and IN 3 is especially difficult for the -R students.

Making Comparisons (IN 1): Grade 8

Easy IN 1 items differ from more difficult IN 1 items in several ways. First, easier items tend to have reading passages that are restricted or

convergent. That is, the elements being compared are clearly presented and clearly suggest to the readers the nature of the comparison they will have to make (see Example 2 next page). Harder IN 1 items have passages that are unrestricted or divergent. That is, the passage does not clearly suggest the nature of the comparison the readers will have to make; only in reading the item stem do the readers learn the nature of the comparison they must make (see Example 1).

Example 1

> Changes occur constantly in space. One of the changes that takes place in space that is most common involves stars. Stars begin during the stage when dust clouds form, called the "birth stage." Then the particles in the cloud exert gravitational pull on each other. Soon the dust cloud contracts, and it gets warm enough to cause a nuclear reaction. This stage is called the "nuclear stage." The hydrogen in the core of the star is converted to helium, and the temperature becomes too low for the nuclear reaction to take place. During this stage, called the "burn-out stage," the star expands and glows red and is called a red giant. Finally, in the "nova stage," the star matter collapses and the star goes from blue to white. The star heats up so much that it reaches temperatures of billions of degrees. The entire life cycle of the star takes billions of years until it becomes a white dwarf.

During which stage of the life cycle of a star is it probably the hottest?

(A) the birth stage

(B) the nuclear stage

(C) the burn-out stage

**(D) the nova stage*

While reading this passage, it is difficult for students, especially -R students, to infer that "hotness" will be the basis for comparing stages in the life cycle of stars.

In general, easier IN 1 items are associated with highly concrete objects for comparison, as well as concrete dimensions of comparison (see Example 2). However, more difficult IN 1 items can be concrete in these ways, but in these items the objects of comparison are typically categorized in superordinate classifications such as "stages," "phases", etc. (see Example 1).

Finally, easier IN 1 Items tend to have wrong options that are negated or contradicted (explicitly or implicitly) in the passage.

Example 2

> Many people like to grow house plants. People like orchids because they have very pretty flowers. Orchids require great care since the soil they are grown in must be kept continuously moist. Ferns grow quite well in certain parts of the house. Ferns like a lot of water and humid air. People who grow jade plants can be lazy because these plants can go for long periods of time without water. Many people grow begonias because begonias grow quickly if they are watered two or three times a week.

1. *Which kind of plant has to be watered least often?*

   (A) begonia

   (B) fern

   **(C) jade*

   (D) orchid

2. *Which kind of plant requires the least attention?*

(A) begonia

(B) fern

**(C) jade*

(D) orchid

Note that in the two items above, the first item involves less inference and is concrete with respect to the objects (plants) and dimension of comparison (watering). The second item above is concrete with regard to objects, but more abstract (therefore requiring more inference) with respect to the comparison dimension of "attention."

Determining Cause and Effect (IN 2): Grade 8

The cause-effect (IN 2) items at grade 8 have three broad levels of difficulty which represent three variations of cause-effect items. The description of these three general types of IN 2 items is designed to illustrate the range of possible items. There is not a large number of each type.

The easiest IN 2 item is markedly easier than the other items. The stimulus material for this item is highly familiar to students and many students could possibly answer the question correctly without reading the passage. The following is an illustration of this type of item.

Example 3

> In earlier times people decorated their Christmas tree with small candles attached to the tree branches. The candles had to be lit each night, watched carefully, and blown out or snuffed out before the candle burned down to the tree. The Christmas trees decorated this way were very pretty but were also very dangerous because the flames from the candles could catch onto the tree and burn down a person's tree and home.
>
> Electric Christmas tree lights have replaced the traditional Christmas candles almost completely. These lights came in many colors and can twinkle and flash on and off. A great value of electric Christmas tree lights is that they are unlikely to start fires.

What is the most likely reason that electric lights have replaced candles as Christmas tree decorations?

ans: Electric lights are safer.

A more difficult type of IN 2 item is clearly more passage dependent. These moderately difficult IN 2 items have the following salient characteristics:

1) They require at least a literal comprehension of the stimulus material.

2) The "cause-effect" relationship in these moderately difficult items is literally stated or requires a very low level of inference to identify.

3) They have a popular incorrect answer which is likely to be generally true in the students' own experience.

4) The correct answer can only be determined from passage information.

The most difficult IN 2 items at grade 8 require a variety of reading and cognitive skills. These difficult IN 2 items require inter-paragraph combining of inferences drawn from the separate paragraphs. The "cause-effect" relationship is generally an abstract concept which is experientially unfamiliar to most of the students. A commonly chosen wrong answer for these items is an accurately described detail or event from the stimulus passage.

A number of these most difficult IN 2 items seem to require a particular cognitive understanding. The basic structure of these items involves two separate "events." Each event has a "cause" and "effect." The "causes" for the two events are similar but differ on one salient feature. The "effects" for the two events are similar but differ in intensity or amount.

An example of this type of item is given below.

Example 4

A sociology student wanted to determine what characteristics of public announcements were most effective in communicating to people. He decided to set up a situation that would answer the question. He chose several messages to be read over the intercom at a local airport. The first day each message was read one time very forcefully. The announcer repeated the word "attention" with each different message. As the student expected, most of the people he asked remembered the messages. The next day, the sociology student had the announcer read the same messages more softly. This time each message was read three times. The announcer did not use the word "attention" this time. The people in the airport remembered the messages more completely than they had the day before.

What is most likely the reason that the second type of message reading strategy was more successful than the first one?

*(A) *People react more positively to repetition than to force.*

*(B) *People in airports pay very little attention to messages.*

*(C) *The second message reading strategy used more interesting messages.*

*(D) *More people were in the airport listening to the messages on the second day.*

A schematic diagram of the cognitive structure implied by this item is shown in figure 9.

EVENT	CAUSE	EFFECT
Day 1:	forcefully read message, read once with the word "attention"	Inference: People remember the message less well than on day 2.
Day 2:	softly read message, read repeatedly, without the word "attention"	People remembered the message more completely than on day 1.

Figure 9. Schematic diagram of difficult IN 2 item, Grade 8

This type of item is not simply assessing student's ability to make an inference about which "cause" produces a particular "effect." This item requires students to identify the difference between two causes which explains the difference between the two effects they produce.

Drawing Conclusions (IN 3): Grade 8

IN 3 at grade 8 is among the most difficult of all BSAP subskills for both +R and -R students and is also one of the most discriminating subskills. There is so little variation among the IN 3 items with respect to difficulty, that comparisons of "easier" to "harder" items reveals little. Analysis of IN 3 items, however, does reveal a number of factors that likely contribute to their overall difficulty.

IN 3 items at grade 8 use a common stem throughout:

Which of the following conclusions can most reasonably be drawn from this selection?

The stem itself represents a difficult reading comprehension task. It is both difficult with regard to its vocabulary load and its level of abstractness. Students who do not have experience with this particular question before seeing it on the examination may not understand what the question is asking them to do.

The IN 3 passages are divergent or unrestricted. They contain no superordinate topic sentence. Therefore, they do not suggest the nature of the task that follows. Only by reading the item stem does the student get a notion of the nature of the task required.

The correct answers for IN 3 items are highly superordinated statements that encompass most or all of the passage information and are therefore highly reading dependent. Correct answers for IN 3 items are also broad statements that apply well beyond the context of and information in the passage. Keyed options are, consequently, broadly applicable rules, generalizations, or principles. To correctly identify the keyed conclusion , students must first summarize the salient information in the passage, then identify a far more general statement that parallels the more specific summary. Additionally, all item options tend to be lengthy, thereby increasing their reading load.

Students' response patterns to IN 3 items suggest they do not understand the nature of the subskill and/or are unable to demonstrate the skill in the context of these particular items and passages. Students

tend to choose wrong answers that repeat a detail from the selection, or are plausible statements based on their own prior information rather than the information in the passage.

Predicting Outcomes (IN 4): Grade 8

The fairly small number of items designed to assess Predicting Outcomes (IN 4) at grade 8 makes it difficult to offer broad generalizations about factors which affect the difficulty of these items. In addition, these items are fairly homogeneous in terms of item difficulty and discrimination. Nonetheless, a number of features differentiating the relatively easy and hard items were noted.

The stimulus passages for the relatively easy items are convergent or restrictive. That is, the stimulus frames a situation or series of events in such a way as to narrow the range of plausible "next events" so that the next event will be an event of a particular kind.

In addition, the relative easy items have an outcome which is the "effect" in a cause-effect relationship. The stimulus passage clearly describes the outcomes (or effects) caused by certain events (or causes). The relationship is established with a structure that can be called "conditional cause-effect". This structure is of the form: if A happens, X follows; if B happens, Y follows; and if C happens, Z follows. The passage makes it clear explicitly or implicitly that A, or B or C has happened and students must identify "what most likely happened next." The next event can be predicted fairly easily after reading the passage. The next event must be X, Y or Z conditional on whether A, B or C occured. A relatively low level of inference is required for these items but a certain logic is

required.

The relatively difficult IN 4 items have stimulus material that is divergent in nature. These passages do not restrict the range of possible "next events" and so the reader must generate a wide range of possibilities while reading the passage. There are no "cause-effect" connections in these items to help the reader. In answering these items, it is very difficult to predict the next event from the passage. The question must be read as, "From among the possible events listed below, what will most likely to happen next?" In at least one case, the event that clearly happens next is not among the options. The keyed correct response may be inconsistent with students' experiences, although it is the most likely outcome of those events that comprise the option set.

Summary

Inference: Grades 6, 8

IN 1: Making Comparisons

IN 2: Determining Cause and Effect

IN 3: Drawing Conclusions

IN 4: Predicting Outcomes

At grade 6, the IN subskills are relatively difficult for both +R and -R students.

Easier IN 1 items and stimulus passages are, in general, more concrete than the harder IN 1 items and passages. Easier IN 1 items require comparisons of elements that have concrete attributes, explicitly communicate

the dimension of comparison, and are associated with highly restricted reading selections. Easier IN 1 items tend to deal with a "best" procedure, and possess wrong answers that are explicitly negated in the selection. More difficult IN 1 items possess less restricted reading selections that also tend to present more information that is irrelevant to the IN 1 task, and require increased inter-sentence combining of information.

Easier IN 2 items often ask for very specific effects for very specific causes that are given in the item passage, while more difficult IN 2 items treat more general causal relationships. Easier IN 2 items occasionally require little more than paraphrasing of text material, and often possess wrong answers that are explicitly negated or contradicted by the stimulus selection. More difficult IN 2 items tend to require more inter-sentence combining of information. The -R students tend to choose wrong answers to IN 2 items that are true in their own experience; -R students often incorrectly infer causal relationships among events simply because they are described in contiguous passage sentences.

All IN 3 items tend to be exceedingly difficult for both +R and -R students. IN 3 items, in general, possess a heavier reading load than other IN subskills. IN 3 items tend to require students to identify salient passage events and description, then superordinate this material into the form of a generalization or summary. Inter-paragraph comprehension is at a premium in IN 3 items.

Easier IN 4 items tend to be associated with reading selections that present an explicit series of actions taken by a single story character, and ask for the "most likely" next event. More difficult IN 4 items are associated with reading passages that possess more complex story and plot lines in which there are multiple significant characters; salient character

information that can significantly influence future events; and a series of actions/events that is less restricted, possibly resulting in any one of a number of future events.

At grade 8, the IN subskills range from difficult to relatively easy for the +R students, and range from moderately to very difficult for the -R students.

Easier IN 1 items are typically accompanied by relatively restricted reading selections, and possess more concrete objects for comparison as well as more concrete dimensions for comparison. The more difficult IN 1 items tend to be accompanied by less restricted reading passages, and often have more abstract objects and dimensions for comparison. The most difficult IN 1 items have reading selections that superordinate comparison object attributes into stages or phases.

Easier IN 2 items appear to be related to reading passages that treat more familiar content. Easier IN 2 items also tend to be answerable more often on the basis of the student's prior information, and generally demand less inferencing. More difficult IN 2 items are associated with reading passages that treat less familiar subjects, and that demand a high degree of inter-paragraph combining of inferences. Harder IN 2 items are also associated with passages that contain multiple causal relationships expressed in the context of a procedural event series (e.g., a research study format).

All IN 3 items tend to be exceedingly difficult. The IN 3 subskill generally places a very heavy reading load on students. The reading selections tend to be highly unrestricted or divergent, and the keyed item option tends to superordinate passage information to a great extent. IN 3 items, consequently, are among the most difficult in BSAP at grade 8.

Easier IN 4 items tend to be associated with reading passages that are highly restricted with regard to the range of possible "next events." Easier IN 4 items tend also to have as keyed options an effect in a cause/effect relationship. More difficult IN 4 items are associated with reading selections that are unrestricted; a range of events could fulfill the criterion of the "most likely" next event. Sometimes in these more difficult IN 4 items, the "most likely" next event is not among the item options, so that the task is one of choosing the "most likely" next event from the ones presented.

ANALYSIS OF LITERATURE

GRADES 6, 8

Subskill Overview

In Analysis of Literature (AL), students are presented reading selections and poems from which they must identify effective uses of language as it communicates feelings and perceptions. They must also critically analyze elements of reading passages and poems by determining accuracy of information and by verifying accuracy through the identification of evidence.

1. Nature of information (AL 1). Students are presented a non-fiction reading passage containing a mixture of facts and opinions, and must distinguish between statements that can be verified by observation (facts) and statements that cannot be verified by observation (opinions) and are only justified with reference to values. Generalizations and inferences are characterized as opinions regardless of their validity since they are not verifiable by observation.

2. Structural Elements (AL 2). Students are presented a brief narrative or dramatic reading passage and must identify the structural elements of plot, character, and setting on the basis of implied cues. At grade 8, the eligible structural elements expand to include mood and point of view.

3. Rhetorical Devices (AL 3). Students are presented a poem that contains multiple rhetorical devices from which they must identify the use of a particular device. At grade 6, eligible rhetorical devices include simile, metaphor, personification and onomatopoeia; at grade 8 the devices eligible for testing expand to include alliteration. The rhetorical device in question is specifically defined in the stem of each question.

100 mm
A5
1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
1.0 mm
1.5 mm
2.0 mm
ABCDEFGHIJKLMNOPQRSTUVWXYZ
abcdefghijklmnopqrstuvwxyz
1234567890

4. Critical Analysis (AL 4). Beginning at grade 8, students are presented a brief narrative or dramatic reading passage which implies relationships among the structural elements (including plot, character, setting, theme, point of view, and mood) it contains. Students must determine the relationship between the structural elements.

Grade 6: Analysis of Literature

Analysis of Subskills 1, 2, and 3

The average proportion of grade 6 students who correctly answer Analysis of Literature items is shown in Table 32 for the students who reach and do not reach grade level standard, along with subskill discriminations.

AVERAGE ITEM DIFFICULTY AND ITEM DICRIMINATION FOR SIXTH GRADE +R AND -R STUDENTS ON ANALYSIS OF LITERATURE SUBSKILLS

TABLE 32

Subskill	Average Difficulty for +R Students	Average Difficulty for -R Students	Average Item Discrimination
AL 1	.81	.34	.47
AL 2	.64	.33	.31
AL 3	.68	.28	.40

The AL subskills are among the most difficult for both the +R and -R students. AL 2 and AL 3 are, in fact, the two most difficult subskills for

the +R group. AL 1 has a discrimination of .46, making it and IN 3 the most discriminating subskills of all for sixth graders.

Nature of Information (AL 1): Grade 6

In grade 6, a relative small number of items assess AL 1. The easy items ask students to identify an opinion from a set of four contiguous sentences in the reading selection. The more difficult items generally ask students to identify a fact.

The nature of the incorrect answers suggests a plausible explanation for the relative difficulties of "opinion" and "fact" question. The easy questions, asking for opinions, have one sentence from the selection that has a signal word like "ought," "should," or signal phrase like "too much" or "too little." The other sentences which constitute incorrect answers describe facts and do not contain these cues.

The more difficult AL 1 items, generally ask students to identify a fact. For these items, one of the incorrect answers contains an opinion signal word like "ought" or "should." Incorrect answers with these signal words are not often selected. The remaining two incorrect answers do not contain opinion signal words. These sentences express generalizations which may be summary statements, inferences, conclusions or predictions. These generalizations may be clearly valid, but nonetheless they are not facts. Students confuse these plausible (and perhaps valid) general statements with facts and frequently select them, incorrectly, as facts.

There is one exception to the general observation that AL 1 items asking for opinions are easy and AL 1 items asking for facts are difficult. A relatively difficult AL 1 item asking students to identify an

opinion seems to interact heavily with students' experiences. This type of item can be illustrated by the following sentence:

Example 1

The NCAA book of track and field records reveals some extraordinary performances by college athletes.

Students would incorrectly identify this sentence as reporting a fact. The sentence contains a true statement but it is not a factual statement in the strictest sense of a fact being verifiable by observation. "Extraordinary" is a value judgment expressing an opinion.

Structural Elements (AL 2): Grade 6

Structural Elements is the most difficult subskill for the +R students (.64) and the third most difficult subskill for the -R students (.33). For these items, students read a passage and then answer questions like, "What does the story tell us about Jeanne?" (e.g., character question) or "What is the story about?" (plot question). Items assessing theme and character appear relatively easier than items measuring plot and setting. These "relatively" easy items are not easy, however. These items are extraordinarily discriminating (greater than .50). On the character items, the -R students are drawn to an incorrect answer which is a specific detail.

Students' responses to the relatively difficult plot and setting items

indicate considerable unfamiliarity with these structural elements. For the -R students, there is one or sometimes two incorrect answers which draw more students than the correct answer. For the plot items, the students select the character option far more often than they select the correct plot answer. On the setting items, a large proportion of the -R students also select the character option.

The high level of difficulty of the AL 2 items and the choices of incorrect answers suggests that a high proportion of grade 6 students are generally unfamiliar with the attributes which define and differentiate the various structural elements.

Rhetorical Devices (AL 3): Grade 6

Rhetorical Devices (AL 3) is the second most difficult subskill for both the +R and for the -R students. The performance of the -R students (.28) is barely above what might be expected if the students simply guessed at all the questions (.25).

The "relatively" easy AL 3 items deal with simile and personification. The relative low difficulty the simile items is not surprising since the stem for all AL 3 items defines the particular rhetorical device in question. The stem for simile questions is:

> *A simile uses the word "like" or "as" to compare two things that are different in most ways. Which of the following lines from the poem contains a simile?*

Students do not actually need to read the poems to answer such questions. They merely need to find the option containing the word "like"

or "as." One of the incorrect options for these items contains the word "as" used as a temporal adverb, e.g., "as I was walking down the street." This option is clearly the most popular incorrect answer. The relative low difficulty of the simile items, and the common choice of the incorrect answer with "as" used as a temporal adverb, suggests that students are using a simple word matching strategy to answer these questions.

The relative low difficulty of the personification items also seems related to the definition of the rhetorical device in the stem. For personification, the stem reads,

> *Personification means describing anything that is not human as if it were human. Which of the following lines in the poem contains an example of personification?*

The option structure for these items generally includes only one option in which the subject of the sentence is inanimate. The definition of personification along with this option structure seems to make these items relatively easy.

Metaphor items are very difficult for both the +R and -R students. In these items, metaphor is defined as:

> *A metaphor compares two things that are different in most ways, without using the words "like" or "as." Which of the following lines from the poem contains a metaphor?*

The option structure for such items generally includes among the options two choices of the forms:

1) The moon is round.

2) The moon is a ball. (Key)

The -R students select type 1 options about as often as type 2 options. There are no signal words to help them with these questions and the absence of signal words increases the difficulty.

All students have a great deal of difficulty with questions about onomatopoeia. Onomatopoeia is defined as:

> *Onomatopoeia means the use of a word that sounds similiar to what it means. Which of the following lines from the poem contains onomatopoeia?*

Many students, it seems reasonable to suggest, do not know how to answer questions involving onomatopoeia. The stem suggests the strategy of thinking about how each word is actually pronounced and checking to see if any pronunciations actually sound like the noise they signify. For example, if students actually thought about the pronounciation of the words in the sentence, "The bacon sizzled on the frying pan," they might recognize that "sizzled" sounds like the noise it signifies. Students' extremely poor performance on these items suggests that they neither understand the concept of onomatopoeia as they read the item nor are they able to gain enough direction from the items to correctly learn the skill as they attempt the item. A final possibility is that they do not know how to pronounce the target words, a deficiency which makes onomatopoeia items almost impossible.

Grade 8: Analysis of Literature

Analysis of Subskills 1, 2, 3, and 4

The average proportion of grade 8 students who correctly answer Analysis of Literature items is shown in Table 33 for the students who reach and do not reach grade level standard, along with subskill discriminations.

AVERAGE ITEM DIFFICULTY AND ITEM DISCRIMINATION FOR EIGHTH GRADE +R and -R STUDENTS ON ANALYSIS OF LITERATURE SUBSKILLS

TABLE 33

Subskill	Average Difficulty for +R Students	Average Difficulty for -R Students	Average Item Discrimination
AL 1	.83	.42	.41
AL 2	.85	.46	.39
AL 3	.80	.40	.40
AL 4	.79	.33	.46

The AL subskills are relatively difficult for the +R students and quite difficult for the -R students. All four subskills are highly discriminating, indicating that they sharply differentiate the two groups.

Nature of Information (AL 1): Grade 8

In grade 8, there is a relatively small number of items assessing AL 1. For the most part, the difficult AL 1 items ask students to identify facts.

The only relatively easy AL 1 item at grade 8 asks for an opinion.

The difficulty of AL 1 at grade 8 appears to be largely a function of students' (1) lack of understanding of what constitutes a "fact," and (2) their subsequent confusion of "facts" with various types of valid inferences. In the example below, this apparent confusion is demonstrated:

Example 2

> [5] Therefore, teachers should be concerned with certification examinations. [6] The Certification Department has made five rulings on appeals this year. [7] The rulings in each instance probably need to be questioned. [8] Teachers must work closely with community groups to keep the certification process from becoming a nightmare.

Look at sentences 5-8 above. Which one states a fact?

(A) Sentence 5

**(B) Sentence 6*

(C) Sentence 7

(D) Sentence 8

In the partial passage above, Sentence 6 is the statement of "fact" that constitutes the correct answer to the item stem. However, each of the other sentences that constitute the incorrect answers can easily be con-

fused as "fact," although each represents an opinion, according to BSAP item structure in AL 1. Students apparently accept as "facts" those statements which are generalizations, conclusions, and probable outcomes so long as these statements are valid. The wrong options in the example item above represent valid inferences and conclusions. "Facts" in this context tend to resemble "details" expressing numerical amounts or events that have occurred. Signaling words like "should," "need," and "probably," that tend to explicitly cue a statement of opinion are apparently disregarded by students when they appear in statements that are valid inferences and conclusions.

Structural Elements (AL 2): Grade 8

The items assessing AL 2 at grade 8 are quite homogeneous within the +R and -R groups. The AL 2 subskill sharply differentiates the +R and -R groups with a subskill discrimination of .39.

The relatively easy items deal with character and mood. The correct answers to these types of items seem to stand out more because there are wrong answers that clearly contradict information from the stimulus material. The pattern of wrong answers does not suggest that the structural elements of character and mood are being confused with any other particular structural elements.

The more difficult AL 2 items all deal with theme or plot. Students' choice of incorrect answer makes it clear that they are confusing these two structural elements. The most popular incorrect answer to plot questions is an option stating the theme of the passage; the most popular incorrect answer to theme questions is an option stating the plot of the

passage.

Both plot and theme options are summary statements with plot options summarizing events and theme options summarizing the overall meaning of passages. The summary nature of plot and theme options appears salient to students but differences in the nature of the passage information being summarized are not clear to them.

Rhetorical Devices (AL 3): Grade 8

Rhetorical devices are difficult for both the +R and -R students at grade 8. The relatively easy AL 3 items at grade 8 deal with similes and alliteration. The presence of signal words or cueing words might explain the relative easiness of these items. The simile items are explicitly defined as including the words "like" or "as." For these items, options with the word "as" used as a temporal adverb are the only plausible distractors if students are using a simple matching strategy.

The AL 3 items assessing alliteration define alliteration as:

> *Alliteration means using several words near each other that begin with the same sound. Which of the following lines from the poem contains alliteration?*

The typical option structure for these items include only one option, the correct answer, in which adjacent words repeat the same sound. This structure makes it very easy to identify "several words near each other that begin with the same sound," e.g., "He snipped the thin thread from the seam."

The more difficult AL 3 item types require students to identify examples of personification and metaphor. The option structure for per-

sonification items includes incorrect answers which are similes and metaphors. The option structure for metaphors generally includes wrong answers which are instances of personification.

Students seem to have only a partial understanding of simile, metaphor and personification. Students are fairly successful at identifying similes by looking for the words "like" or "as." Students appear to know that metaphors involve comparisons, but confuse these comparisons with personifications which "compare" non-human things to humans by giving human attributes to the non-human entities.

Consider the following three sentences.

Simile:	*The monkey stared at the labyrinth like a scholar.*
Metaphor:	*The monkey was a scholar staring at the labyrinth.*
Personification	*The scholarly monkey studied the labyrinth.*

The basic meaning (semantic content) of these three sentences is virtually identical, but the particular form used to express the common meaning leads to different rhetorical classifications for each sentence.

Critical Analysis (AL 4): Grade 8

Critical Analysis (AL 4) items are a type of cause-effect question. Students read a passage, and are then asked to identify something from the story (a cause) that would lead to a certain inference or conclusion about some structural element of the story like mood or character. For example,

What happens in the story to create a mood of horror?
What happens in the story to tell us Jack has a strange sense of humor?

These items are extremely difficult for both the +R and -R students and they are also highly discriminating. Data are available for only a small number of AL 4 items, hence a detailed description is not possible. Questions about character seem to be easier than question about mood. The very low performance of the -R students suggests that they may not understand the nature of the task.

Summary

Analysis of Literature: Grades 6, 8

AL 1: Nature of Information

AL 2: Structural Elements

AL 3: Rhetorical Devices

AL 4: Critical Analysis

The AL subskills at grade 6 range from moderately difficult to very difficult for the +R students, and range from difficult to very difficult for the -R students.

Easier AL 1 items ask grade 6 students to identify an opinion from among a set of contiguous sentences in the selection. Correct answers that are opinions are explicitly cued as such by words like "should," "ought," and etc. and so forth. More difficult AL 1 items ask for facts, and are

uncued by signal words. More difficult AL 1 items tend to possess popular incorrect answers that express valid inferences, but that are not facts. Occasionally, students' performance on a particular item is clearly affected in a large way by their prior information regarding the topic treated in the selection.

Easier AL 2 items tend to focus on the structural elements of theme and character, while the more difficult AL 2 items deal with elements of plot and setting. The easier and the harder AL 2 items tend to have very popular incorrect options. Response patterns of students suggest that they are highly unfamiliar with literary structural elements, and/or the particular circumstances in which structural elements are tested.

Easier grade 6 AL 3 items focus on simile and personification. More difficult AL 3 items assess students' understanding of metaphor and onomatopoeia. Identification of simile and personification is likely easier than metaphor and onomatopoeia due to explicit cueing of the presence of these devices. Metaphor and onomatopoeia remain relatively uncued. Students' response patterns suggest a high degree of unfamiliarity with metaphor and onomatopoeia.

At grade 8, the AL subskills are moderately difficult for the +R students, and range from difficult to very difficult for the -R students.

Easier AL 1 items at grade 8 ask students to identify an opinion from among a set of four contiguous sentences repeated from the selection. Correct answers that are opinions are explicitly cued as such by words like "should," "ought," and so on. More difficult AL 1 items ask for facts, and are uncued by signal words. More difficult AL 1 items tend to possess popular answers that express valid conclusions, inferences, or other types of valid generalizations. Students appear to confuse the truth/validity of

a statement with whether or not the statement is factual.

The easier AL 2 items at grade 8 focus on the structural elements of character and mood, while the more difficult AL 2 items focus on theme and plot. Students appear to confuse theme with plot and plot with theme in the more difficult AL 2 items.

The easier AL 3 items focus on grade 8 students' understanding of simile and alliteration. The more difficult AL 3 items focus on students' understanding of personification and metaphor. The difficulty of these respective skills is largely attributable to the extent to which the skills are explicitly cued, both in the reading selections and in the items. Students' response patterns suggest they are confusing personification with metaphor.

Finally, the AL 4 items are all extremely difficult for both groups of students. Items about characters appear to be somewhat easier than items about mood. The relatively small number of AL 4 items makes it difficult to observe any additional response patterns of students.

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

This study was designed to answer four major questions. These questions are:

(1) Have students in South Carolina made progress toward achieving the grade level standards in reading and mathematics from 1981 to 1985?

(2) Do the students who achieve both the reading and mathematics standards in 1985 reach the reading standard by as much as the corresponding students in 1981? And, are the students who reach neither the reading nor the mathematics standards in 1985 any closer to reaching the reading standard than the corresponding students in 1981?

(3) Are there specific areas of reading in which students across the state have deficiencies? If so, what are the deficiency areas?

(4) What actions might be taken to improve students' achievement relative to the BSAP reading instructional objectives based on the information from questions 1, 2 and 3?

This section of the report, organized around the four research questions, summarizes the research findings and provides final conclusions regarding the research questions. A summary of the research findings related to questions 1, 2 and 3 are presented followed by a discussion of how the research results can be used to improve students' achievement in those areas of reading in which students still have difficulty.

Tables 2 and 8 and Figure 7 reported earlier in the text are repeated in this section for the convenience of the reader. Readers are urged to remember that material presented in this section is a summary of a more comprehensive presentation of the same material and that additional information can be obtained by examining the appropriate sections of the complete report. It may be particularly helpful for readers to note that a complete description of each reading objective and reading subskill is contained in the body of this report. This description can be used as a "mini T&T."

Summary of Questions 1, 2 and 3

Research Question 1: *Have students in South Carolina made progress toward achieving the grade level standards in reading and mathematics from 1981 to 1985?*

The answer to this question is clearly "yes". Table 2 on the next page indicates that for all grades studied there has been an average increase of 16.1% in the percentage of students who met both standards in 1981. The highest increase is in grade 3 with a gain of 19.6%. Students in grades 6 and 8 showed the lowest gain with a 12.3% and 13.3%, respectively. Table 2 also shows an average decrease of 13.5% in the percentage of students who met neither standard. All five grades studied were close to this average.

Combining the increased percentage of students reaching both standards and the decreased percentage of students who reach neither standard provides an index of the shift in total achievement of the students

Number (and Percentage) of Students Meeting or Failing to Meet
BSAP Standards in Reading and Mathematics

Grade 1	All Ss	+R+M	-R+M	+R-M	-R-M
1985	49,782	37,754 (75.8%)	3,806 (7.7%)	2,830 (5.7%)	5,392 (10.8%)
1984	49,572	36,304 (73.2%)	3,937 (8.0%)	3,364 (6.8%)	5,967 (12.0%)
1983	46,440	31,207 (67.2%)	4,015 (8.7%)	3,782 (8.1%)	7,436 (16.0%)
1982	46,516	28,076 (60.4%)	3,379 (7.3%)	5,409 (11.6%)	9,652 (20.8%)
1981	47,412	28,245 (59.6%)	4,190 (8.8%)	4,908 (10.4%)	10,069 (21.2%)
Grade 2	**All Ss**	**+R+M**	**-R+M**	**+R-M**	**-R-M**
1985	44,944	33,076 (73.6%)	5,356 (11.9%)	2,501 (5.6%)	4,011 (8.9%)
1984	42,590	29,698 (69.7%)	5,415 (12.7%)	2,469 (5.8%)	5,008 (11.8%)
1983	43,239	27,312 (63.2%)	5,552 (12.8%)	3,124 (7.2%)	7,251 (16.8%)
1982	44,226	24,641 (55.7%)	3,865 (8.7%)	5,660 (12.8%)	10,060 (22.8%)
1981	45,491	24,728 (54.4%)	6,865 (15.1%)	3,273 (7.2%)	10,620 (23.3%)
Grade 3	**All Ss**	**+R+M**	**-R+M**	**+R-M**	**-R-M**
1985	42,219	30,513 (72.3%)	2,952 (7.0%)	4,855 (11.5%)	3,899 (9.2%)
1984	42,840	29,402 (68.6%)	4,447 (10.4%)	3,554 (8.3%)	5,437 (12.7%)
1983	43,957	28,246 (64.3%)	4,465 (10.2%)	4,939 (11.2%)	6,307 (14.4%)
1982	45,207	25,555 (56.5%)	5,062 (11.2%)	5,624 (12.4%)	8,966 (19.8%)
1981	46,952	24,743 (52.7%)	4,065 (8.7%)	6,952 (14.8%)	11,192 (23.8%)
Grade 6	**All Ss**	**+R+M**	**-R+M**	**+R-M**	**-R-M**
1985	44,923	23,647 (52.6%)	3,949 (8.8%)	6,309 (14.0%)	11,018 (24.5%)
1984	46,515	23,015 (49.5%)	3,433 (7.4%)	7,176 (15.4%)	12,891 (27.7%)
1983	49,987	24,093 (48.2%)	4,117 (8.2%)	6,458 (12.9%)	15,319 (30.7%)
1982	48,127	21,827 (45.4%)	2,526 (5.3%)	8,190 (17.0%)	15,584 (32.4%)
1981	46,546	18,764 (40.3%)	3,278 (7.0%)	6,832 (14.7%)	17,672 (38.0%)
Grade 8	**All Ss**	**+R+M**	**-R+M**	**+R-M**	**-R-M**
1985	48,955	24,370 (49.8%)	3,723 (7.6%)	6,998 (14.3%)	13,864 (28.3%)
1984	47,135	21,530 (45.7%)	3,867 (8.2%)	6,807 (14.4%)	14,931 (31.7%)
1983	45,969	16,868 (36.7%)	2,468 (5.4%)	9,032 (19.7%)	17,601 (38.2%)
1982	45,504	16,033 (35.2%)	2,452 (5.4%)	7,744 (17.0%)	19,275 (42.4%)
1981	46,667	17,024 (36.5%)	2,901 (6.2%)	6,795 (14.6%)	19,947 (42.7%)

Code

+R+M: Student Passing Both Tests
-R+M: Students Failing Reading, Passing Mathematics
+R-M: Students Passing Reading, Failing Mathematics
-R-M: Students Failing Both Tests

at each grade. This shift ranges from 25.9% at grade 6 to 34.2% at grade 3, with an average of 29.6% for all grades.

Research Question 2: *Do the students who achieve both the reading and mathematics standards in 1985 reach the reading standard by as much as the corresponding students in 1981? And, are the students who reach neither the reading nor the mathematics standards in 1985 any closer to reaching the reading standard than the corresponding students in 1981?*

The first part of this question arose out of a concern that instruction aimed at helping more students reach the grade level standard might detract from instruction aimed at challenging students who could easily reach the standard and who should, therefore, be achieving at a level well above the end-of-year standard. The second part of this question was designed to determine if the students most in need of help were moving closer to the grade level standard in reading even if they had not yet reached that goal.

This question was examined by classifying students according to how far their performance was from the grade level standard. The students of interest were those who were "well above standard" or "well below standard." "Well above standard" was defined as scoring above the standard by more than two standard errors of measurement (S.E.M.). "Well below standard" was defined as scoring below the standard by more than two standard errors of measurement. The results of this analysis are summarized in Table 8 on the next page.

1981-1985 Changes in the Percentage of Students
Passing or Failing the Reading Standard
by Greater than two S.E.M.'s

TABLE 8

	1981-1985 Change in the Percentage of +R+M Students Meeting the Standard by Greater Than 2 S.E.M.'s			1981-1985 Change in the Percentage of -R-M Students Failing the Standard by Greater Than 2 S.E.M.'s		
Grade	1981	1985	Change	1981	1985	Change
1	81.9	85.9	4.0	61.8	61.2	- 0.6
2	76.5	87.7	11.2	84.3	83.2	- 1.1
3	75.8	80.5	4.7	74.2	64.0	-10.2
6	72.0	73.1	1.1	70.9	60.9	-10.0
8	68.4	69.0	0.6	74.5	66.1	- 8.4
Average	74.9	79.2	+4.3	73.1	67.1	- 6.0

The information in Table 8 indicates that there is an increase in the percentage of of +R+M students (those who reach both the reading and mathematics standard) passing the reading standard by greater than two S.E.M.'s in all grades. In the -R-M group (those who reach neither standard) there is a decrease in the percentage of students failing the reading standard by greater than two S.E.M.'s in all grades.

These data indicate that the answer to question 2 for grades 1, 2, 3, 6, and 8 is clearly "yes," although improvements in the +R+M group for grade 8 and in the -R-M group in grade 1 are not especially strong.

Research Question 3: *Are there specific areas of reading in which students across the state have deficiencies? If so, what are the deficiency areas?*

Despite clear overall progress in student achievement from 1981 to 1985, there were specific areas of reading in which students had deficiencies. These deficiencies were studied at the objective, subskill, and item levels. The proportion of students correctly answering each test item was the smallest unit of analysis. This proportion is called the item difficulty. The subskill difficulty is the average item difficulty for all items measuring a subskill; the difficulty of an objective is the average item difficulty of all items measuring an objective. Item subskill and objective discrimination refers to the extent to which an item, subskill or objective differentiates between those who reach grade level standard and those who do not reach the standard.

Student Performance on the Reading Objectives

The six BSAP reading objectives are Decoding and Word Meaning, Details, Main Idea, Reference Usage, Inference, and Analysis of Literature. The pattern of student achievement across these objectives in grades 1, 2, 3, 6, and 8 is quite systematic. The six objectives cluster into two groups of three objectives, one of which is relatively easy and one of which is relatively difficult. The relatively easy objectives are Reference Usage (RE), Details (DE), and Decoding and Word Meaning (DW); the more difficult objectives are Inference (IN), Main Idea (MI), and Analysis of Literature (AL). Student performance on the objectives can be shown by graphing the average item difficulty for each objective for the +R students and for the -R students. Figure 7 displays this information for sixth grade students on the next page. The patterns seen for grade six are similar to the patterns in the other grades although the pattern is most obvious at grade 6.

The left vertical line in Figure 7 represents the performance of the +R students and right vertical line represents performance of the -R students. For both groups of students, the RE-DE-DW cluster is clearly easier than the IN-MI-AL cluster. The difference between the two clusters is more pronounced for the -R students.

The easiest objective for the -R students is RE. On average, 65.1% of the -R students correctly answer RE items. On the other hand, the most difficult objective for the +R students is AL. On average, 70.9% of the +R students answer AL items correctly. In comparing these findings it is

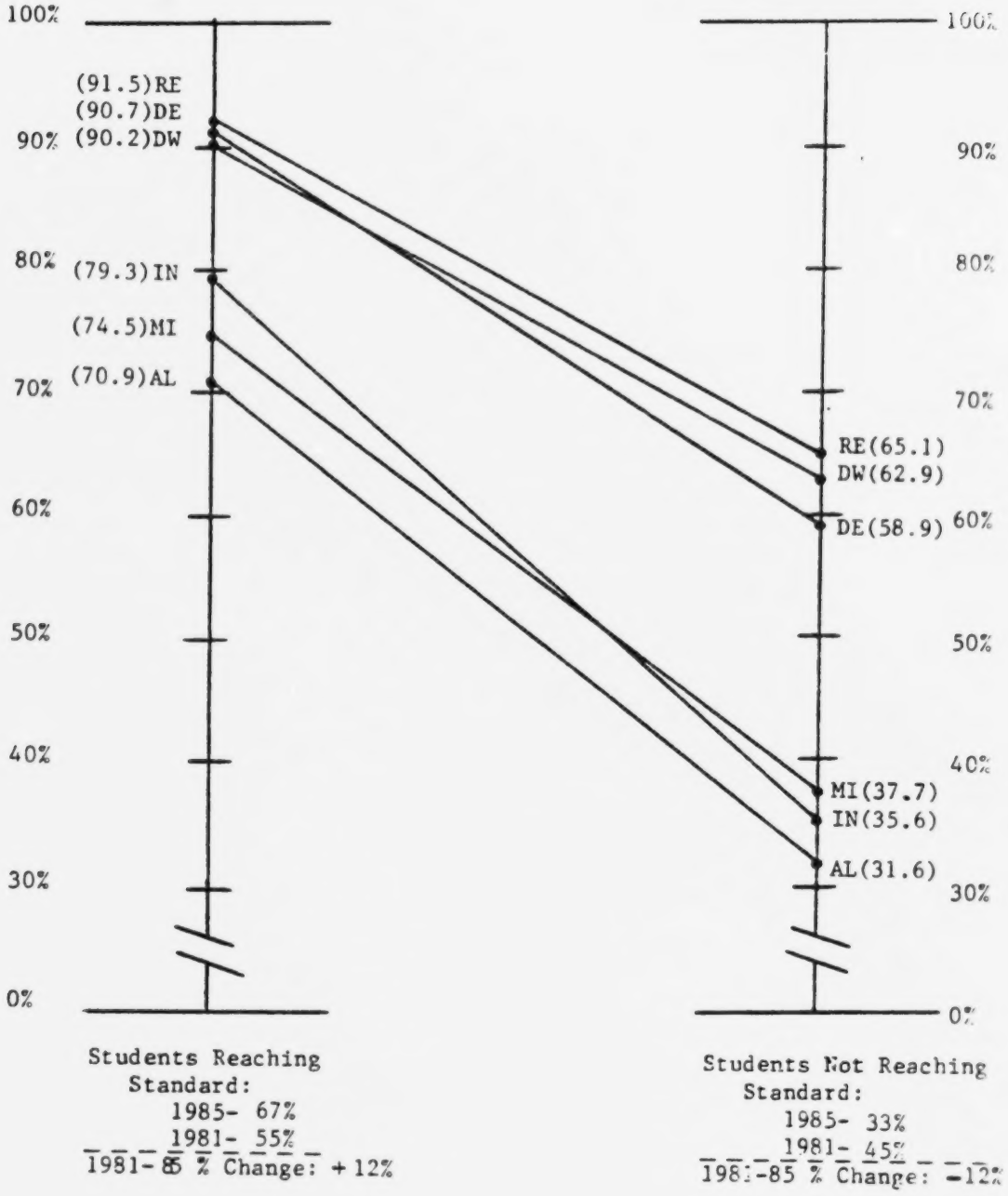


Figure 7. : Average percentage of Grade 6 students reaching and not reaching standard who correctly answer items in the six BSAP objectives

clear that the best performance of the -R students (RE) is 5% below the worst performance of the +R students. Thus the BSAP objectives are very sharply discriminating between the +R and -R students. Similar trends are observed at grades 1, 2, 3 and 8.

Student Performance at the Subskill Level

Students' performance on BSAP subskills is examined by analyzing the individual test items assessing each subskill and by averaging the item difficulties for all items measuring each subskill. Throughout the subskill analyses, comparisons of the +R and -R students are performed. The Reading Subskills within the six BSAP objectives are as follow:

DW: Decoding and Word Meaning

- DW 1: Sight Recognition (tested at all grades)
- DW 2: Phonetic Decoding (tested at grades 1, 2, 3)
- DW 3: Contextual Word Meaning (tested at grades 2, 3, 6, 8)
- DW 4: Structural Word Meaning (tested at grades 6, 8)

DE: Details

No subskills (tested at all grades)

MI: Main Idea

- MI 1: Restated Main Idea (tested at grades 1, 2, 3)
- MI 2: Paraphrased Main Idea (tested at all grades)
- MI 3: Inferred Main Idea (tested at all grades)

RE: Reference Usage

- RE 1: Selecting a Reference Source (tested at all grades)
- RE 2: Using a Reference Source (tested at all grades)

IN: Inference

- IN 1: Making Comparisons (tested at all grades)

IN 2: Determining Cause and Effect (tested at all grades)
IN 3: Drawing Conclusions (tested at all grades)
IN 4: Predicting Outcomes (tested at all grades)

AL: Analysis of Literature

AL 1: Nature of Information (tested at all grades)
AL 2: Structural Elements (tested at all grades)
AL 3: Rhetorical Devices (tested at all grades)
AL 4: Critical Analysis (tested at grade 8)

DECODING AND WORD MEANING: GRADES 1, 2, 3

DW 1: Sight Recognition

DW 2: Phonetic Decoding

DW 3: Contextual Word Meaning

DW 1: Sight Recognition

The +R students across grades 1 through 3 perform very well in Sight Recognition, while the -R students perform relatively well in DW 1 at grades 1 and 2, but lower at grade 3. Sight Recognition items tend to differentiate the +R and -R students very well at all grades.

Performance differences between +R and -R students in DW 1 are mainly due to the amount of experience students have had with the vocabulary being tested. While the target vocabulary words are at the appropriate level with respect to the S.C. Word List, difficult items tend to focus on words that likely appear less frequently in grade level instructional material.

DW 2: Phonetic Decoding

The +R students perform extremely well in Phonetic Decoding at grades 1 through 3. The -R students have their lowest performance at grade 1. The -R students' performance at grades 2 and 3 is similar to that of +R students. Phonetic Decoding items tend to discriminate between +R and -R students only moderately well at grade 1, and not at all in grades 2 and 3.

DW 3: Contextual Word Meaning

DW 3 is not assessed at grade 1. The +R students perform moderately well at grade 3 on Contextual Word Meaning items. The -R students at grade 2 demonstrate low performance on DW 3 items. The -R students perform moderately well at grade 3 on DW 3 items. Contextual Word Meaning items at grades 2 and 3 tend to differentiate very well between +R and -R students.

The more difficult DW 3 items in grades 2 and 3 appear to possess an increased reading load that requires more inferential processing. This contrasts to the easier DW 3 items in which the relevant contextual information is typically contained in a single word or word pair, is explicit in nature, and/or directly cued by means of key word and signal word repetition between stimulus and item stem.

DETAILS: GRADES 1, 2, 3

No Subskills

GRADE 1

Items assessing Details are moderately difficult for the +R students and extremely difficult for the -R students. The -R students do not seem to understand the task required by these items. The responses of these -R students suggest they do not understand that they are supposed to answer the questions based on information contained in the stimulus selection. They often seem to answer based on their own prior information and general knowledge. Questions requiring students to identify "when" an event occurred are particularly difficult for first graders.

GRADE 2

In grade 2 the Details questions are quite easy for the +R students and moderately difficult for the -R students. Items based on reading selections that introduce two or more characters and refer to story characters with pronouns are most difficult. When pronoun referencing of story characters extends across two or more sentences, the items are especially difficult.

GRADE 3

At grade 3 the Details items are very easy for the +R students and moderately difficult for the -R students. "Sequence-of-events" questions

that do not repeat from the stimulus signal word cues such as "first," "after," "last," etc., in the correct answer are clearly more difficult than questions which repeat these signal word cues.

MAIN IDEA: GRADES 1, 2, 3

MI 1: Restated Main Idea

MI 2: Paraphrased Main Idea

MI 3: Inferred Main Idea

The MI items are relatively easy or very easy for the +R students and uniformly difficult or very difficult for the -R students. The MI items sharply differentiate the +R from the -R group at grades 1, 2, and 3.

GRADE 1

Restated Main Idea items with the main idea appearing at the beginning of the stimulus passage are relatively easy. The -R students often select incorrect options which are true statements in the childrens' experience, even if unrelated to the stimulus and task. The difficulty of Paraphrased Main Idea items is affected by whether the main idea is directly and explicitly stated or indirectly and implicitly stated. Paraphrased Main Idea questions in which the main idea is supported by several or all of the sentences tend to be easier. In the Inferred Main Idea items, the -R students frequently choose wrong options heavily cued with simple word or phrase repetitions from the stimulus. Difficult items repeat words or phrases from the reading selection in the incorrect answers. The -R

students frequently select options which are true in their general experience even if unrelated to the stimulus and not responsive to the task.

GRADE 2

The difficulty of Restated Main Idea items depends on the clarity with which the main idea stands out in the stimulus material. Test items in which the main idea is a clearly stated topic sentence, heavily supported by the other sentences, are easier than other types of items. The relatively easy Paraphrased Main Idea items also have main ideas that are clearly stated topic sentences. Incorrect options which repeat verbatim details are commonly selected by -R students. Difficulty of Inferred Main Idea items is strongly affected by the number of sentences from which students must draw inferences and the amount of inter-sentence information combining required. Easier MI 3 items have a main idea that can be inferred from a single major topic sentence alone.

GRADE 3

Restated Main Idea items with the main idea in the first or last sentence of the stimulus material are easier than other items. Items with simple and clearly stated main ideas, supported explicitly by all the other sentences, are relatively easy. Paraphrased Main Idea items with well cued, explicitly supported topic sentence main ideas are easier than other types of MI 2 items. The -R students frequently pick options which repeat a detail from the selection if the main idea is not simple and

obvious. Inferred Main Idea items, as with most of the main idea items, vary in difficulty depending on the degree to which the main idea is implied. Easier items and selections have main ideas that can be inferred from just one or two sentences of the stimulus passage.

REFERENCE USAGE: GRADES 1, 2, 3

RE 1: Selecting a Reference Source

RE 2: Using a Reference Source

The Reference Usage subskills at grades 1 through 3 are among the easiest in BSAP. The +R students at each grade perform very well in RE 1 and RE 2, while the -R students at each grade perform moderately well for the most part. Reference Usage items tend to discriminate well between +R and -R students in grades 1 through 3.

GRADES 1 AND 2

At grades 1 and 2, the difficulty of RE 1 and RE 2 items is primarily related to the particular reference sources used as stimulus material. The -R students at grades 1 and 2 appear to be unfamiliar with the nature of picture dictionaries and tables of contents (RE 1). The -R students at grade 2 also have difficulty with RE 2 items that require them to get information from picture dictionaries, but perform relatively well on RE 2 items that require them to get information from tables of contents.

GRADE 3

At grade 3, the type of reference source strongly influences the difficulty of RE 1 items, but has less influence on RE 2 items. Difficult RE 1 items at grade 3 tend to focus on tables of contents and encyclopedias, while easy items focus on dictionaries and to some extent maps.

The difficulty of RE 2 items at grade 3 is primarily influenced by the complexity--rather than type--of the stimulus reference source and the specific nature of the skill and task demands. In general, items that require (1) locating more information, (2) getting information from several locations in the stimulus source and (3) comprehending a sentence tend to be more difficult.

INFERENCE: GRADES 1, 2, 3

IN 1: Making Comparisons

IN 2: Determining Cause and Effect

IN 3: Drawing Conclusions

IN 4: Predicting Outcomes

At grades 1 through 3, Inference items tend to range from moderately easy to very easy for +R students and difficult to very difficult for -R students. All Inference items tend to sharply differentiate +R from -R students.

GRADE 1

At grade 1, direct and explicit cueing has a large effect on the difficulty of Inference items. In Making Comparisons (IN 1), selections and items that directly name or cue the dimension used for the comparison (e.g., "nearest," "farthest," etc.) tend to be easier. Items are much more difficult when students must infer both the dimension for the comparison and the relative position of elements being compared.

Item difficulty is also affected by the amount of inter-sentence comprehension an item requires. In Predicting Outcomes (IN 4) items, easier items tend to describe a series of events, each of which cues the correct answer. More difficult IN 4 items require students to comprehend multiple events in sequence in order to correctly identify a probably outcome.

Finally, items which include pronoun/antecedent substitutions tend to be more difficult. In Drawing Conclusions (IN 3), for example, stimulus selections that name an agent or character and then subsequently refer to that agent in the selection and item stem with pronoun substitution tend to be more difficult than IN 3 items that do not use pronoun referencing.

GRADE 2

The difficulty of Inference items is affected by explicit cueing between stimulus selection and items via key word and/or signal word repetition. The more difficult items cue incorrect answers, for example in Drawing Conclusions (IN 3). Key vocabulary repetition in wrong options appears to draw students, particularly -R students, to those wrong

wrong answers.

The relative degree of inferencing required to correctly answer Inference items at grade 2 also influences item difficulty. In Predicting Outcomes (IN 4) items, the easier items typically possess incorrect options that are explicitly negated or contradicted in the selection. Harder IN 4 items tend to possess incorrect options that are implicitly contradicted in the accompanying selections.

GRADE 3

The type of cueing available to students influences Inference item difficulty. In Making Comparisons (IN 1), easier items typically explicitly name the dimension of comparison in one word. More difficult IN 1 items tend to use a phrase or clause to suggest the dimension of comparison.

Vocabulary cueing of wrong answers is associated with more difficult items in Drawing Conclusions (IN 3). These items tend to possess incorrect options that repeat key words from the selection and draw more student responses particularly from among the -R students.

Predicting Outcomes (IN 4) items that possess explicitly negated or contradicted incorrect options tend to be easier. Harder items possess incorrect options that are implicitly contradicted in the selection.

ANALYSIS OF LITERATURE: GRADES 1, 2, 3

AL 1: Nature of Information (Make-believe)

AL 2: Structural Elements (Plot, Character, Setting)

AL 3: Rhetorical Devices (Rhyme, Simile)

GRADE 1

Nature of Information (AL 1) items are moderately difficulty for the +R students and extremely difficult for the -R students. The -R students do not understand what is meant by "make-believe" or they do not understand the nature of the test questions. The -R students frequently select options that make sense or are true in their own experience. Students are more often able to answer correctly when the "make-believe" element involves people rather than animals.

In Structural Elements (AL 2), plot and setting items tend to be easier than items dealing with the analysis of characters. The "character" items tend to require more inference and inter-sentence comprehension. Rhetorical Devices (AL 3) items require the identification of pairs of rhyming words from a stimulus poem. This task seems unfamiliar to many students, but particularly the -R students.

GRADE 2

In grade 2 all the AL subskills are extremely difficult for the -R students and quite easy for the +R students. In AL 1, the -R students appear not to understand the concept of "make-believe" and are unfamiliar with the vocabulary words that cue such items. In AL 2, character items

are more difficult than plot and setting items. As in grade 1, character items require more inference and inter-sentence comprehension than plot and setting. AL 3 performance of the -R students suggests that they are unfamiliar with the concept "rhyming-words," or they do not know the pronunciation of vocabulary words in the items needed to correctly identify the rhyming words.

GRADE 3

In grade 3, AL 1 and AL 2 subskills are very easy for the +R students and relatively easy for the -R students. AL 1 items involving a mythical character are more difficult than other types of AL 1 items. Difficult AL 2 items tend to require inter-sentence comprehension and contain little explicit cueing. Both the +R and -R students have great difficulty identifying members of a simile comparison in the AL 3 items. The extremely low performance and the choice of options suggests that students do not have a clear understanding of the skill and task demands of items that require analysis of poems to identify elements being compared.

GENERAL SUMMARY OF READING SUBSKILL PERFORMANCE: GRADES 1, 2, 3

Following are a number of general observations regarding student performance on BSAP Reading subskills in Grades 1, 2 and 3. These observations are a result of the analyses previously described, and are included here since they describe test item and stimulus selection

characteristics that tend to appear consistently across Reading subskills in these early grades.

Literal Comprehension and Inferential Comprehension

Student performance on the reading objectives reveals a cluster of relatively easy objective and a cluster of relatively difficult objectives. The easy objectives are Reference Usage (RE), Details (DE), and Decoding and Word Meaning (DW). These easy objectives measure relatively low level literal comprehension. The difficult objectives are Inference (IN), Main Idea (MI), and Analysis of Literature (AL). The difficult objectives measure relatively complex inferential comprehension. All students find the inferential comprehension objectives more difficult than the literal comprehension objectives. All the objectives differentiate students who reach grade level standard from those who do not. The inferential comprehension objectives differentiate the two groups very sharply.

General Task Familiarity

Item response analyses suggest a general task familiarity issue. Students, especially -R students, may not fully understand that information must be derived from reading selections in answering BSAP Reading items. Students apparently do not fully understand that prior knowledge and general information about the world is not sufficient for answering items. Students may not understand the more specific tasks inherent in the various Reading items; i.e., "making comparisons," "determining cause and effect," and so on.

Key and Signal Word Cueing

The presence or absence of explicit cueing (i.e., the item stem repeats key or signal words from the stimulus passage) has a large effect on how well students comprehend text. When correct answers are cued in this way, the items are much easier for students. The -R students tend to rely very heavily on direct and explicit cueing and tend to answer items incorrectly when incorrect options are cued.

Pronoun Substitution

A special type of cueing involves pronoun substitution for characters and agents being asked about in the items. When pronoun substitutions are used on reading selections, "who" kinds of items become much more difficult. When pronoun substitutions extend across sentences in the reading selection, and when the item stems themselves use pronoun substitution, the difficulty of the reading task is compounded greatly, particularly for -R students.

Inter-Sentence Comprehension

Items in many subskills require students to "use" information from several sentences rather than from a single sentence. Regardless of the subskill, this characteristic makes reading tasks more difficult. This particular characteristic extends to single word reading versus phrase and clause reading, suggesting that the larger the unit of words students must read and understand to answer items, the more difficult the items will be.

DECODING AND WORD MEANING: GRADES 6,8

DW 1: Sight Recognition

DW 3: Contextual Word Meaning

DW 4: Structural Word Meaning

GRADE 6

DW 1 and DW 3 items appear relatively easy for both +R and -R students; both groups, however, demonstrate significantly lower performance on DW 4 items.

Easier DW 1 items target vocabulary words that likely appear more frequently in reading materials, and are therefore more familiar than the words that appear in the harder DW 1 items. Easier DW 3 items tend to be heavily cued by single words in context. More difficult DW 3 items cue the meaning of the unfamiliar target word with a phrase, thus increasing the reading and comprehension load. More difficult DW 3 items also employ negated phrase cueing where the item asks for the unnegated target word meaning, and students must reason through the task using their understanding of opposition in meaning. Easier DW 4 items tend to have target words that possess more familiar root word forms, while more difficult items possess more unfamiliar root word forms. More difficult DW 4 items also tend to possess target words with prefixes that are less restricted (e.g., "con") in their meaning than the easier DW 4 items which contain words with more restricted prefixes (e.g., "pre").

GRADE 8

The DW 1 and DW 4 items are relatively difficult for both the +R and -R students, while DW 3 items are relatively easy for both groups.

DW 1 represents the most difficult of the Decoding and Word Meaning subskills for grade 8 students. Easier items tend to focus on vocabulary that likely appear more frequently in reading materials, while more difficult DW 1 items focus on words that appear less frequently.

DW 3 items at grade 8 are highly homogeneous, revealing little pattern in student performance. However, DW 4 items that are easier tend also to focus on more familiar vocabulary and vocabulary possessing prefixes with highly restricted meanings. The more difficult DW 4 items, therefore, focus on less familiar root forms that tend to possess prefixes with variable and less restricted meanings. Difficult DW 4 items also occasionally focus on derived word forms using suffixes that change syntactic and grammatic function of the target words.

DETAILS: GRADES 6, 8

No Subskills

GRADE 6

At grade 6, the items measuring Details are very easy for the +R students and moderately difficult for the -R students.

Easier DE items at grade 6 ask for a specific piece of information or an event contained in a sequence of events. These items typically repeat verbatim in the correct answer the necessary information from the stimulus passage, and in the instance of events in a sequence, repeat a signal word from the passage.

More difficult DE items at grade 6 possess little of the key and signal word repetition that characterizes the easier DE items. Very difficult DE items require students to identify an event in a sequence that occurred "just before" or "just after" some other event; items of this type possess no cueing through vocabulary word repetition.

GRADE 8

At grade 8, items measuring Details are easy for the +R students, and moderately difficult for the -R students.

Easier DE items at grade 8 tend to possess cueing between item and stimulus passage in the form of vocabulary repetition. More difficult DE items possess little or no such cueing and they focus on target sentences that contain two events or elements, only one of which is relevant to the question posed in the item. Finally, DE items that ask for events "just before" other events are exceedingly difficult.

MAIN IDEA: GRADES 6, 8

MI 2: Paraphrased Main Idea

MI 3: Inferred Main Idea

GRADE 6

The MI subskills at grade 6 are among the most difficult in BSAP for this grade level for both +R and -R students.

Easier MI 2 items tend to have clear and identifiable topic sentences that are only slightly different in wording from the correct answer. More difficult MI 2 items tend to have one or more of the following

characteristics in common: a two paragraph stimulus passage with the second paragraph containing the necessary information for identifying the paraphrased main idea; salient main idea information expressed in the context of idiomatic language; a stimulus passage with no topic sentence that requires students to identify salient information/events; and/or highly popular incorrect options.

Easier MI 3 items and reading selections have several common features. In these easier items, the stimulus selection may not possess a topic sentence, but does provide a great deal of concrete support for the main idea, often through repetition of salient information. Easier MI 3 items also tend to have as correct answers what amounts to a summary of the selection. More difficult MI 3 items are associated with reading selections that possess a particular construction. In these passages, two paragraphs that are causally linked form the selection. Identifying the Inferred Main Idea in this type of passage requires students to link salient information from the two paragraphs, thus increasing the inter-paragraph reading demands of these passages and items. Difficult MI 3 items also tend to have very popular wrong answers.

GRADE 8

The MI subskills at grade 8 are moderately easy, relatively speaking, for both the +R and -R students.

Easier MI 2 items at grade 8 have several features in common. The topic of the related reading passage in these items tends to be highly familiar and correct answers tend to be word-for-word paraphrases of topic sentences from the passage. Also, easier MI 2 items accompany passages

that explicitly cue or mark the location of the topic sentence for the reader.

More difficult MI 2 items and their associated passages have one or more of the following features in common: passages may present a sequence of events in mixed order and posses far less cueing of topic sentences; passages often contain vocabulary words that are both unfamiliar and salient to identifying the main idea; correct answers are often phrase or clause paraphrases of passage information, rather than word-for-word paraphrases; and often the MI 2 task requires the student to combine information from multiple topic sentences which have "subordinate/superordinate" relationship.

Easier MI 3 items at grade 8 tend to require students either to make inferences from just one paragraph or short segment of the stimulus reading passage, or to combine literal information from across paragraphs. More difficult MI 3 items tend to require a greater degree of inference, the combining of inferred information from across paragraphs of the reading passage, and possess one or more incorrect answers that are very appealing to students.

REFERENCE USAGE: GRADES 6, 8

RE 1: Selecting a Reference Source

RE 2: Using a Reference Source

GRADE 6

Relatively speaking, RE subskills are among the easiest for both +R and -R students at grade 6. In RE 1, students appear to have a firm

understanding of the types of information contained in telephone directories, dictionaries, and encyclopedias. On the other hand, they exhibit confusion with the types of information contained in indexes, card catalogs, and classified ads; this is particularly true for the -R students.

Easy RE 2 item can be answered by matching a key vocabulary word from the item stem to the same word in the stimulus, then locating the required information. More difficult RE 2 items cannot be answered via this matching strategy. RE 2 items that require students to sort through different levels of subordination/superordination as typically exist in a book index tend to be more difficult.

GRADE 8

In RE 1 items, grade 8 students appear to be familiar with the type of information contained in telephone directories, card catalogs, and dictionaries. However, grade 8 students, particularly -R students, demonstrate confusion with the kinds of information contained in book bibliographies, newspaper indexes, and classified ads. The relative plausibility of wrong answers greatly influences item difficulty in RE 1 items. Students, and particularly -R students, demonstrate some confusion in differentiating between specific reference sources (e.g., a newspaper index) and more general terms that might be considered equivalent (e.g., chart since a newspaper index is a type of chart).

INFERENCE: GRADES 6, 8

IN 1: Making Comparisons

IN 2: Determining Cause and Effect

IN 3: Drawing Conclusions

IN 4: Predicting Outcomes

GRADE 6

At grade 6, the IN subskills are relatively difficult for both +R and -R students.

Easier IN 1 items and reading selections are, in general, more concrete than the harder IN 1 items and passages. Easier IN 1 items require comparisons of elements that have concrete attributes, explicitly communicate the dimension of comparison, and are associated with highly restricted reading selections. Also, easier IN 1 items tend to require students to identify a "best" procedure, and possess wrong answers that are explicitly negated in the selection. More difficult IN 1 items possess less restricted reading selections, present information that is irrelevant to the IN 1 task, and require increased inter-sentence combining of information.

Easier IN 2 items often ask for very specific effects for very specific causes that are given in the item stem, while more difficult IN 2 items treat more general causal relationships. Easier IN 2 items occasionally require little more than paraphrasing of text material, and often possess wrong answers that are explicitly negated or contradicted by the stimulus selection. More difficult IN 2 items tend to require more

inter-sentence combining of information. The -R students tend to choose wrong answers to IN 2 items that are true in their own experience, as well as to incorrectly infer causal relationships among events that are described in contiguous passage sentences.

All IN 3 items tend to be exceedingly difficult for both +R and -R students. IN 3 items, in general, possess a heavier reading load than other IN items. IN 3 items tend to require students to identify salient passage events and description, then superiordinate this material into the form of a broad and valid generalization. Inter-paragraph comprehension is critical in IN 3 items.

Easier IN 4 items tend to be associated with reading selections that present an explicit series of actions taken by a single story character, and ask for the "most likely" next event. More difficult IN 4 items are associated with reading passages that possess more complex story and plot lines in which there are several significant characters; salient character information that can significantly influence future events; and a series of actions/events that is less restricted, possibly resulting in any one of a number of future events.

GRADE 8

At grade 8, the IN subskills range from difficult to relatively easy for the +R students, and range from moderately difficult to very difficult for the -R students.

Easier IN 1 items are typically accompanied by relatively restricted reading selections, and possess more concrete objects for comparison as

well as more concrete dimensions for comparison. The more difficult IN 1 items tend to be accompanied by less restricted reading passages, and often have more abstract objects and dimensions for comparison. The most difficult IN 1 items have reading selections that superordinate comparison object attributes into stages or phases.

Easier IN 2 items appear to be related to reading passages that treat more familiar content. Easier IN 2 items also tend to be answerable more often on the basis of the student's prior information, and generally demand less inferencing. More difficult IN 2 items are associated with reading passages that treat less familiar subjects and demand a high degree of inter-paragraph combining of inferences. Harder IN 2 items are also associated with passages that contain multiple causal relationships expressed in the context of a series of steps in a procedure.

All IN 3 items tend to be exceedingly difficult. The IN 3 subskill generally places a very heavy reading load on students. The reading selections tend to be highly unrestricted or divergent, and the correct answers tend to superordinate passage information into broad generalizations. IN 3 items, consequently, are among the most difficult in BSAP at grade 8.

Easier IN 4 items tend to be associated with reading passages that are highly restricted with regard to the range of possible "next events." Easier IN 4 items also tend to have as correct answers an effect in a cause/effect relationship. More difficult IN 4 items are associated with reading selections that are unrestricted; a range of events could fulfill the criterion of the "most likely" next event. Often in these more

difficult IN 4 items, the "most likely" next event is not among the item options, so that the task is one of choosing the "most likely" next event from among those presented.

ANALYSIS OF LITERATURE: GRADES 6, 8

AL 1: Nature of Information

AL 2: Structural Elements

AL 3: Rhetorical Devices

AL 4: Critical Analysis

GRADE 6

The AL subskills at grade 6 range from moderately difficult to very difficult for the +R students, and range from difficult to very difficult for the -R students.

Easier AL 1 items ask students to identify an opinion from among a set of contiguous sentences from the selection. Correct answers that are opinions are explicitly cued as such by signal words like "should," "ought," and so on. More difficult AL 1 items tend to possess popular incorrect answers that express valid inferences but are not facts. Occasionally, students' performance on a particular item is clearly affected in a large way by their prior information regarding the topic treated in the selection.

Easier AL 2 items tend to focus on the structural elements of theme and character, while the more difficult AL 2 items deal with elements of plot and setting. The easier and the harder AL 2 items tend to have very

popular incorrect options. Response patterns of students suggest that they are highly unfamiliar with literary structural elements, and/or the particular circumstances in which Structural Elements are tested.

Easier AL 3 items focus on similes and personification. More difficult AL 3 items assess students' understanding of metaphor and onomatopoeia. Identification of similes and personification is likely easier than metaphor and onomatopoeia due to explicit cueing of the presence of these devices. Metaphor and onomatopoeia remain relatively uncued. Students' response patterns suggest a high degree of unfamiliarity with metaphor and onomatopoeia. It is important to note that the definitions of the Rhetorical Devices are given with the test items.

GRADE 8

The AL subskills are moderately difficult for the +R students, and range from difficult to very difficult for the -R students.

Easier AL 1 items ask students to identify an opinion from among a set of four contiguous sentences repeated from the selection. Correct answers that are opinions are explicitly cued as such by signal words as "should," "ought," and so on. More difficult AL 1 items ask for the identification of facts, and are uncued by signal word repetition. More difficult AL 1 items tend to possess popular answers that express valid conclusions, inferences, or other types of valid generalizations. Students appear to confuse the truth/validity of a statement with whether or not the statement is factual.

The easier AL 2 items focus on the structural elements of character

and mood, while the more difficult AL 2 items focus on theme and plot. Students appear to confuse theme with plot and plot with theme in the more difficult AL 2 items.

The easier AL 3 items focus on students' understanding of similes and alliteration. The more difficult AL 3 items focus on students' understanding of personification and metaphor. The difficulty of these respective skills is largely attributable to the extent to which the skills are explicitly cued, both in the reading selections and in the items. Students' response patterns suggest they are confusing personification with metaphor.

Finally, the AL 4 items are all extremely difficult for all groups of students. Items about characters appear to be somewhat easier than items about mood.

GENERAL SUMMARY OF READING SUBSKILL PERFORMANCE: GRADES 6, 8

Following are several general observations regarding the performance of sixth and eighth grade students on BSAP Reading subskills. These observations emanate from the analyses previously described, and are included here since they describe test item, stimulus selection, and task demand characteristics that tend to appear consistently across Reading subskills in grades 6 and 8.

Literal Comprehension and Inferential Comprehension

Student performance on the reading objectives reveals a cluster of relatively easy objective and a cluster of relatively difficult objectives. The easy objectives are Reference Usage (RE), Details (DE) Decoding and Word Meaning (DW). These easy objectives measure relatively low level literal comprehension. The difficult objectives are Inference (IN), Main Idea (MI), and Analysis of Literature (AL). The difficult objectives measure relatively complex inferential comprehension. All students find the inferential comprehension objectives more difficult than the literal comprehension objectives. All the objectives differentiate students who reach grade level standard from those who do not. The inferential comprehension objectives differentiate the two groups very sharply.

Key and Signal Word Cueing

The presence or absence of explicit vocabulary cueing (e.g., the item stem/options repeat key or signal word from the selection) has a large effect on how well students comprehend text. When correct answers are explicitly cued in this way, the items are much easier for students. Clearly, the -R students tend to rely very heavily on the presence of direct and explicit cueing of this sort, suggesting a "search-and-match-the-word" strategy for reading comprehension. Students, especially -R students, have a strong tendency to choose wrong answers if the answers are cued by key or signal word repetition.

Inter-Paragraph Combining of Information/Inferences

In literal or inferential reading comprehension items and tasks, the more explicit information students must draw together from the selection in order to correctly answer an item, the more difficult is the item and task. Similarly, the more inferences students must make and then combine in order to complete an additional inference task, the more difficult is the corresponding item and task. Generally, the more widely distributed the task-relevant information in a reading selection, the more the item and task load on "inter-paragraph" reading and the more difficult is the corresponding item and task.

Subordinate/Superordinate Relationships

Common to a variety of literal and inferential reading subskills is the necessity of recognizing and understanding subordinate and superordinate text relationships. These relationships exist in text in many ways, and range from very concrete (e.g., cat, dog, and cow can be superordinated into the category "animals") to far more abstract text relations. Any item or task that requires students to superordinate text information into broader and more inclusive categories will tend to be very difficult. This appears to be the case even in recognition tasks where students need only to recognize that superordination has taken place, and are not required to engage in superordinating objects or ideas into more inclusive categories. The more abstract the superordinate categories and levels are, the more difficult the task.

Restricted/Unrestricted Reading Selections

Highly restricted or convergent reading selections which focus on a single topic or theme tend to be easier to read than unrestricted or divergent reading selections, regardless of the type of subskill being assessed. Restricted reading selections tend to be highly suggestive of the question or reading subskill task that is to follow. Unrestricted reading selections tend to be more difficult since they suggest a far broader range of possible questions or tasks that would likely follow. Simple story structures containing concrete events tend to be highly restricted, while more complex story structures that contain a wider array or relevant information and events tend to be less restricted. More restricted reading passages tend to converge on a more restricted range of possible questions, while unrestricted passages tend to be more divergent with respect to follow-up tasks and questions.

Restricted reading selections tend also to be more conventionally structured (e.g., the topic sentence is typically the first sentence of the selection) than unrestricted selections. The more conventional patterning of the reading selection may well contribute to a related observation: that regardless of reading subskill being assessed, students tend to have great difficulty identifying salient and important passage information. The degree of restrictedness of a reading selection likely contributes to students' ability to identify important passage information, which in turn underlies performance in Main Idea and Inference subskills.

Research Question #4:

In responding to this question, it is useful to enumerate the areas in which the research shows students have deficiencies. In grades 1, 2, and 3, students exhibit general problems with those Reading objectives which required inferential rather than literal comprehension skills. Specific subskill deficiency areas were described in summary fashion in the preceding section. Four general problem areas across the various subskills where students have consistent difficulties include task familiarity, vocabulary cueing, pronoun substitution, and general inter-sentence reading comprehension.

As related to issues of teacher and classroom practice, the results of the present investigation suggest the efficacy of (1) allocating increased instructional time to the deficiency areas noted, and (2) increasing the amount of direct instruction in the deficiency areas noted.

The body of this report describes and provides examples of many features of written text that appear to relate to student comprehension difficulties. At grades 1 through 3, some of the specific areas that appear to warrant increased allocation of instructional time and provision of more direct reading comprehension instruction include helping students with:

1. The general concept of getting explicit information from text.
2. The ability to read and comprehend a variety of anaphoric text relations at the sentence level, including pronoun substitution.

3. The ability to use negated and situational context structures to gain meaning for unfamiliar vocabulary.
4. The general ability to connect information from across sentences, or inter-sentence comprehension.
5. The general concepts of Main Idea and Inference.
6. The general ability to identify important information in text, and to recognize (and create) summaries of that information.
7. The ability to recognize topic sentences and to differentiate supporting from non-supporting sentences.

In grades 6 and 8, students also demonstrate difficulties with those BSAP Reading objectives that require inferential comprehension, rather than literal comprehension. Specific subskill deficiency areas were described in summary fashion in the preceding section. Four general problem areas across the various subskills where students have consistent difficulties include vocabulary cueing, inter-paragraph inferencing and combining of information, recognition and comprehension of subordinate-superordinate relationships, and comprehension of unrestricted and less highly structured reading selections.

As related to issues of teacher and classroom practice, the results of the present investigation suggest the efficacy of (1) allocating increased instructional time to the deficiency areas noted, and (2) increasing the amount of direct instruction in the deficiency areas noted. The present study also reports numerous examples of how specific features of written text appear to relate to student comprehension difficulties at grades 6 and 8. Some of the more basic areas that seem to warrant direct reading comprehension instruction include:

1. Recognition of phrase, clause, and sentence paraphrases that are syntactically altered (from the way in which they appear in text).
2. General recognition and comprehension of a sequence of events, particularly when the sequence is presented in temporally mixed order in text.
3. General ability to link information from across paragraphs of text, or inter-paragraph reading comprehension.
4. General concept of inferring information from text and how such inference processes differ from getting the main idea.
5. General ability to identify important text information, and recognize when text information is task-irrelevant.

Additionally, this report notes and provides many examples of written text features that further complicate the reading demands placed upon students at grades 6 and 8. Many of these reading and text-processing skills, and are at a premium in reading content area textbooks. They include:

6. Ability to derive the main idea from multiple-paragraph text selections that do not contain clear topic sentences nor conventional paragraph structuring.
7. Ability to recognize the subordinate/superordinate text relations that typically link concept elements in, particularly, content textbook exposition.
8. Ability to recognize how specific text information can support broadly applicable generalizations.
9. General concept of "likelihood" and probability as they affect predicting outcomes in narrative and expository text.

10. Ability to recognize valid and invalid causal connections among events in a sequence.

11. Ability to differentiate between factual statements and valid inferences.

It is difficult for any one group of educators to decide exactly how to respond to the findings of this research. A useful general approach, however, can be suggested. This general approach involves the use of study groups composed of educators and other interested persons. Such groups can be established and used at the State, district, or school level. Study groups at all these levels would surely be appropriate. It is critical that such study groups represent a wide range of people interested in and knowledgeable about reading education. Such study groups should not be limited to school personnel alone.

Study groups should examine this report and based on their own analyses combined with their own experiences, the study groups should make recommendations in response to the following questions:

1) What strategies, procedures and materials can be obtained or developed to support the reading curriculum?

2) What strategies, procedures and materials can be obtained or developed to support reading instruction?

3) What strategies, procedures and materials can be obtained or developed to support reading teachers? (The response to this question would address both preservice and inservice teacher training needs.)

While study and policy recommending groups might well make a large number of very different recommendations in response to these four

questions given the needs of the students, the teachers, and the schools involved, the authors of this research report recommend that all policy groups consider two of the basic recommendations already made in this section. The first has to do with reconsidering the use of classroom time:

(1) Instructional time should be allocated to working with students in the deficiency areas noted.

The second has to do more with how to utilize more effectively the classroom time that is available for instruction:

(2) Direct instruction should be given students in the deficiency areas noted.

The authors of the present investigation recognize that study groups will, and should, vary in how they deal with implementing changes relative to the findings of this study. The authors, however, firmly believe that more effective allocating of classroom instructional time to teaching important and basic reading skills in a more direct fashion underly the successful implementation of these research findings.